GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - nucleic search, using frame_plus_p2n model

Run on: March 26, 2004, 17:14:33 ; Search time 4233 Seconds
                                            (without alignments)
                                 2856.770 Million cell updates/sec

Title: US-10-014-927-19
Perfect score: 1451
Sequence: 1 MSSRWNRTIVGNLPGDIRK.........RSKSRSRSRSNSPVSPVISG 279

Scoring table: BLOSUM62
                Xgapop 10.0 , Xgapext   0.5
                Ygapop 10.0 , Ygapext   0.5
                Fgapop 6.0 , Fgapext   7.0
                Delop   6.0 , Delext   7.0

Searched: 3470272 seqs, 2167151699S residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US10014927/runat_24032004_152925_9896/app_query.fasta_1.455
 -Db=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US10014927 @CGN_1_3731@runat_24032004_152925_9896 -NCPU=6 -ICPU=3
 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   GenEmbl:*
           1:   gb_ba:*
           2:   gb_htg:*
           3:   gb_in:*
           4:   gb_om:*
           5:   gb_ov:*
           6:   gb_pat:*
           7:   gb_ph:*
           8:   gb_pl:*
           9:   gb_pr:*
          10:   gb_ro:*
          11:   gb_sts:*
          12:   gb_sy:*
          13:   gb_un:*
          14:   gb_vi:*
          15:   em_ba:*
          16:   em_fun:*
          17:   em_hum:*
          18:   em_in:*
          19:   em_mu:*
          20:   em_om:*
          21:   em_or:*
          22:   em_ov:*
          23:   em_pat:*
          24:   em_ph:*
          25:   em_pl:*
          26:   em_ro:*
          27:   em_sts:*
          28:   em_un:*
          29:   em_vi:*
          30:   em_htg_hum:*
          31:   em_htg_inv:*
          32:   em_htg_other:*
          33:   em_htg_mus:*
          34:   em_htg_pln:*
          35:   em_htg_rod:*
          36:   em_htg_mam:*
          37:   em_htg_vrt:*
          38:   em_sy:*
          39:   em_htgo_hum:*
          40:   em_htgo_mus:*
          41:   em_htgo_other:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1377.5	94.9	838	8	AY150486	AY150486 Arabidops
2	1272	87.7	762	6	AX506504	AX506504 Sequence
3	1032.5	71.2	933	8	AY050912	AY050912 Arabidops
4	904	62.3	59261	8	T12M4	AC003114 Arabidops
5	891	61.4	5164	8	ATH131214	AJ131214 Arabidops
6	875	60.3	4044	6	AX040661	AX040661 Sequence
7	865	59.6	1229	8	AK118379	AK118379 Arabidops
8	865	59.6	1232	8	AY085920	AY085920 Arabidops
9	854.5	58.9	1523	8	AY056185	AY056185 Arabidops
10	854.5	58.9	2059	8	AK118074	AK118074 Arabidops
11	847	58.4	858	8	BT006316	BT006316 Arabidops
12	847	58.4	1540	8	AY128338	AY128338 Arabidops
13	846	58.3	1293	8	AK071503	AK071503 Oryza sat
14	839	57.8	1160	8	ATHRIBNUCP	M98340 Arabidopis
15	832.5	57.4	1352	8	AK106176	AK106176 Oryza sat
16	827.5	57.0	1238	8	AK061903	AK061903 Oryza sat
17	827.5	57.0	3430	8	AK121333	AK121333 Oryza sat
18	816	56.2	1899	8	AK072839	AK072839 Oryza sat
19	808	55.7	1501	8	AK065547	AK065547 Oryza sat
20	757	52.2	2196	8	AK103534	AK103534 Oryza sat
21	586.5	40.4	1618	5	BC042354	BC042354 Xenopus l
22	585	40.3	2369	5	BC046679	BC046679 Xenopus l
23	584.5	40.3	2765	6	BD189938	BD189938 Tumor ant
24	582.5	40.1	1428	6	BD094070	BD094070 Shear str
25	582.5	40.1	1428	9	HUMSF2P33	M69040 Human SF2p3
26	582.5	40.1	2708	9	BC010264	BC010264 Homo sapi
27	582.5	40.1	2765	9	AB062124	AB062124 Homo sapi
28	582.5	40.1	2785	10	BC058627	BC058627 Mus muscu
29	582.5	40.1	2878	10	BC046773	BC046773 Mus muscu
30	566.5	39.0	809	3	AF242767	AF242767 Caenorhab
31	566	39.0	2002	5	BC055511	BC055511 Danio rer
32	551.5	38.0	94536	8	T14P8	AF069298 Arabidops
33	551.5	38.0	194892	8	ATCHRIV6	AL161494 Arabidops
34	550	37.9	1099	3	AK115991	AK115991 Ciona int
35	550	37.9	1391	3	AF232773	AF232773 Drosophil
36	550	37.9	1395	3	AY051919	AY051919 Drosophil
37	546.5	37.7	160415	9	AC140961	AC140961 Papio anu
38	543	37.4	1584	6	AX305513	AX305513 Sequence
39	543	37.4	1584	10	MMASF	X66091 M.musculus
40	542.5	37.4	1717	9	HUMASF	M72709 Human alter
41	542.5	37.4	2926	8	AF173640	AF173640 Arabidops
42	541.5	37.3	5285	6	AR204400	AR204400 Sequence
43	539.5	37.2	1860	5	BC056752	BC056752 Danio rer
44	525	36.2	1193	10	BC012217	BC012217 Mus muscu
45	520	35.8	257703	3	CEY11B2A	AL132904 Caenorhab

ALIGNMENTS

RESULT 1

```
AY150486
LOCUS       AY150486                838 bp    mRNA    linear   PLN 23-SEP-2002
DEFINITION  Arabidopsis thaliana putative SF2/ASF splicing modulator Srp30
            (At1g09140) mRNA, complete cds.
ACCESSION   AY150486
VERSION     AY150486.1  GI:23297698
KEYWORDS    FLI_cDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
            Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 838)
  AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
            Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
            Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     Annotation is based on the January 2002 version of the Arabidopsis
            genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1..838
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="C105028"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51."
     gene            1..838
                     /gene="At1g09140"
     CDS             1..807
                     /gene="At1g09140"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative SF2/ASF splicing modulator Srp30"
                     /protein_id="AAN13011.1"
                     /db_xref="GI:23297699"
                     /translation="MSSRWNRTIYVGNLPGDIRKCBVEDLFYKYGPIVDIDLKIPPRP
                     PGYAFVEPEPRDADDAIYGRDGYDFDGCRLRVEIAHGGRRFSPSVDRYSSYSASRA
                     PSRRSDYRVLVTGLPPSASWQDLKDHMRKAGVCFSEVFPDRKGMSGVVDYSNYDDMK
                     YAIRKLDATEFRNSAYIRVREYESRSVSRSPDDSKSYSRSRSRGPSCSYSSKSR
                     SVSPARSISPRSRPLSRSRSLYSSVSRSQSRSKSRSRSRSNSPVSPVISG"
     3'UTR           808..838
                     /gene="At1g09140"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.09e-66        Length:         838
Score:                  1377.50         Matches:        268
Percent Similarity:     96.06%          Conservative:   0
Best Local Similarity:  96.06%          Mismatches:     0
Query Match:            94.93%          Indels:         11
DB:                                     Gaps:           1
```

```
US-10-014-927-19 (1-279) x AY150486 (1-838)

Qy     1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG 60

Qy    21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    61 TGTGAGGTTGAAGATCTCTTCTACAAGTATGGACCAATTGTGGACATTGATTTGAAGATT 120
                                                                         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    41 ProProArgProProGlyTyrAlaPheValGluProGluProArgAspAlaAspAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCACCGAGACCTCCTGGTTATGCCTTTGTCGAGTTTGAGCCACGATCCTCGTGATGCAGACGAT 180

Qy    61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GCAATTTATGGACGTGATGGTTATGATTTTGATGGGTGTCGACTTCGGGTTGAGATTGCA 240

Qy    81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATGGTGGTCGTAGATTTTCACCATCAGTTGATAGGTACAGCAGCAGCTACAGTGCGAGC 300

Qy   101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
         · ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CGTGCACCTTCAAGACGCTCTGACTACCGGTCTTGTGACCGGATTACCGCCCTTCTGCT 360

Qy   121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyValCysPheSerGluVal 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCGTGGCAGGACCTTAAGGACCATACACATGCGCAAAGCTGGAGATGTCGCTTCTCGAAGTT 420

Qy   141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTCCCTGACCGTAAAGCATGTCTCGGGGTTGGATTATAGCAACTATGATGATGAAG 480

Qy   161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TACGCAATAAGGAAACTTGATGCCACTGAATTTCGAAATGTTTCCTAGTGCTATATA 540

Qy   181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CGGGTGAGGGAATATGAGTCAGAGAGTGTGAGTCGAAGCCCAGATGAATTCAAAAGTAT 600

Qy   201 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGAAGCAGGAGTCGGAGCCGTGGTCCAAGCTGTAGCTATAGTAGCAAGAGCAGGAGTGTG 660

Qy   221 SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCACCTGCTAGATCCATTTCCCGCGTTCACGGCCCCTTAGTCGTTCTCGCTCGCTATAC 720

Qy   241 SerSerValSerArgSerGlnSerArgSerLysSerArgSerArgSerArg 260
           ||||||||||||||||||||||||||||||||||||
Db   721 AGTCTGTCTCAAGG----------------------------TCCCAATCAAGA 747

Qy   261 SerLysSerArgSerArgSerArgSerAsnSerProValSerProValIleSerGly 279
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748 TCAAAATCAAGATCAAGATCAAGATCGAATTCTCCAGTTTCACCTGTGATATCTGGT 804
```

```
RESULT 2
AY506504
LOCUS       AX506504       762 bp   DNA    linear   PAT 27-SEP-2002
DEFINITION  Sequence 1199 from Patent WO0216655.
ACCESSION   AX506504
VERSION     AX506504.1  GI:23387741
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
  TITLE     Stress-regulated genes of plants, transgenic plants containing
            same, and methods of use
  JOURNAL   Patent: WO 0216655-A 1199 28-FEB-2002;
            The Scripps Research Institute (US) ; Syngenta Participations AG
            (CH)
FEATURES             Location/Qualifiers
     source          1..762
                     /organism="Arabidopsis thaliana"
```

```
DEFINITION   Arabidopsis thaliana putative SF2/ASF splicing modulator Srp30
             protein (At1g09140) mRNA, partial cds.
ACCESSION    AY050912
VERSION      AY050912.1  GI:15292956
KEYWORDS     FLI cDNA.
SOURCE       Arabidopsis thaliana (thale cress)
  ORGANISM   Arabidopsis thaliana
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1  (bases 1 to 933)
  AUTHORS    Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
             Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
             Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
             Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
             Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
             Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
             Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE      Arabidopsis Full Length cDNA Clones
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 933)
  AUTHORS    Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M.,
             Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L.,
             Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
             Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
             Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,
             Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M.,
             Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
             Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and
             Theologis,A.
  TITLE      Direct Submission
  JOURNAL    Submitted (24-JUL-2001) Plant Gene Expression Center, 800 Buchanan
             Street, Albany, CA 94710, USA
COMMENT      RIKEN Genomic Sciences Center (GSC) members carried out the
             collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
             Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
             Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
             Hayashizaki,Y. and Shinozaki,K.

             The Salk, Stanford, PGEC (SSP) Consortium members carried out the
             sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X.,
             Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D.,
             Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
             Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H.,
             Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E., Lam,B.,
             Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
             Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.

             Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
             this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
             contributed equally to this work as PIs.

             Annotation is based on the January 2002 version of the Arabidopsis
             genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1..933
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="RAFL08-12-I05 (R11038)"
                     /ecotype="Columbia"
                     /note="This clone is in a modified pBluescript vector
                     (FLC-1) as a BamHI/XhoI insert."
     gene            <1..933
                     /gene="At1g09140"
     CDS             <1..626
                     /gene="At1g09140"
                     /codon_start=3
                     /evidence=experimental
                     /product="putative SF2/ASF splicing modulator Srp30
                     protein"
                     /protein_id="AAK93589.2"
```

ORIGIN   /mol_type="unassigned DNA"
                         /db_xref="taxon:3702"

Alignment Scores:
Pred. No.:   2.17e-60   Length: 762
Score: 1272.00   Matches: 253
Percent Similarity: 91.67%   Conservative: 0
Best Local Similarity: 91.67%   Mismatches: 1
Query Match: 87.66%   Indels: 23
DB: Gaps: 2

US-10-014-927-19 (1-279) x AX506504 (1-762)

```
Qy   1   MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys  20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   ATGAGTAGCCGATGGAATCGTACGATCTACCGTGGAATTTGCCTGGAGATATTCGCAAG   60

Qy   21  CysGluValGluAspLeuPheTyrLysPheTyrGlyProIleValAspLeuLysLysIle  40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  TGTGAGGTTGAAGATCTCTTCACAAGTATGGACCAATTGTCGACATTGATTTGAAGATT  120

Qy   41  ProProArgProProGlyTyrAlaPheValGluHisGluProArgGlyAspAlaAspAsp  60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCACCGGACCTCCTGTTTATGCCTTTGTCGAGTTTGAAGATCCTGTGATGCAGACGAT  180

Qy   61  AlaIleTyrArgGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla  80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GCAATTTATGGACGTGGATGGTTATGATTTGATGGGTGTCGACTTCGGGTTGAGATTGCA  240

Qy   81  HisGlyGlyTyrArgArgPheSerProSerValAspArgArgTyrSerSerArgSerSer  100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CATGGTGGTTATAGAAGATTTTCACCATCAGTTGATAGGTACAGCAGCTACAGTGCCGAGC  300

Qy   101 ArgAlaProSerArgArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CGTGCACTTCAAGACGCTCGACTACCGCCGTGCTTGTGACCGGATTACCGCCTTCTGCT  360

Qy   121 SerTrpGlnAspLeuLysAspHisMetArgGlyAlaGlyValCysPheSerGluVal  140
         |||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCGTGGCAGGACCTTAAGGATCACATGCGCAAAGCTGGAGATGTCTGCTTCTCTGAAGTT  420

Qy   141 PheProAspAspArgLysGlyMetSerGlyValValValAspTyrSerAsnTyrAspAspMetLys  160
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TTCCCTGACCGTAAAGGCATGTCTGGGGTTGTGGATTATAGCAACTATGATGATATGAAG  480

Qy   161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgGlyAsnAlaPheSerSerAlaTyrIle  180
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TACGCAATAAGGAAACTTGATGCCACTGAATTTCGAAATGCTTTCTCCAGTGCTTATATA  540

Qy   181 ArgValAlaArgGluTyrGluSerArgSerValSerArgSerProAspAspSerProLysSerTyr  200
         |||||||||||||||||||||||||||||||||
Db   541 CGGGTGGAGGGAATATGAGTCGAGGAGTGTGAGTCGAGCCCAGATGATTCTAAAAGCTAT  600

Qy   201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal  220
         |||||||||||||
Db   601 AGAAGCAGGAGTCGG----------------------------AGCCGTGT-GTG   626

Qy   221 SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyr  240
         |||||||||||||||||||||||||||||||||||||
Db   627 TCACCTGCTAGATCCATTTCCCGGTTCACGGCCCCTAGTCGTTCTCGCTCGCTATAC  686

Qy   241 SerSerValSerArgSerGlySerLeuLeuArgAlaAlaAspTrpIleSerGlnSerArg  260
         |||||||||||||||||||||||||||||||||
Db   687 AGCTCTGTCTCAAGG-------------------------TCCCAATCAAGA  713

Qy   261 SerLysSerArgSerArgSerArgSerArgSerAsnSerProValSerProVal  276
         |||||||||||||||||||||
Db   714 TCAAAATCAAGATCAAGATCAAGATCGAATTCTCCAGTTTCACCTGTA  761
```

```
RESULT 3
AY050912
LOCUS       AY050912       933 bp    mRNA      linear      PLN 18-SEP-2002
```

```
                                        ||||||||::|||||||||||||||||||||||||||||||||||||
                         Db      570 AAATCAAGAACAAGATCAAGATCGAATTCCCAGTTTCACCTGTGATATCTGGT 623

                         RESULT 4
                         T12M4
                         LOCUS        T12M4              59261 bp    DNA    linear   PLN  23-JUN-1998
                         DEFINITION   Arabidopsis thaliana chromosome 1 BAC T12M4 sequence, complete
                                      sequence.
                         ACCESSION    AC003114
                         VERSION      AC003114.1  GI:3249094
                         KEYWORDS     HTG.
                         SOURCE       Arabidopsis thaliana (thale cress)
                         ORGANISM     Arabidopsis thaliana
                                      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                                      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                                      rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                         REFERENCE    1 (bases 1 to 59261)
                         AUTHORS      Vysotskaia,V.S., Schwartz,J.R., Toriumi,M., Yu,G., Oji,,O.,
                                      Kwan,,A., Liu,S., Li,J., Araujo,R., Au,M., Brendel,V., Buehler,E.,
                                      Conway,A.B., Conway,A.R., Dewar,K., Feng,J., Kim,C., Kurtz,D.,
                                      Li,Y., Palm,C.J., Shinn,P., Sun,H., Davis,R.W., Ecker,J.R.,
                                      Federspiel,N.A. and Theologis,A.
                         TITLE        Arabidopsis thaliana chromosome 1 BAC T12M4 sequence, complete
                                      sequence
                         JOURNAL      Unpublished (1998)
                         REMARK       This sequence is of BAC T12M4 from Arabidopsis thaliana chromosome
                                      1. The sequence does not represent the sequence of the entire
                                      insert of this clone. It is shorter by 11040 bp because we submit
                                      only the unique sequence of the clone. However, in order to
                                      facilitate the joining of overlapping clones in the future for
                                      creation of larger contigs, we provide a small overlap (200 bp)
                                      between overlapping submitted clones. The 3' end of this sequence
                                      overlaps by 200 bp the 5' end of the sequence of the BAC F7G19.
                         REFERENCE    2 (bases 1 to 59261)
                         AUTHORS      Theologis,A.
                         TITLE        Direct Submission
                         JOURNAL      Submitted (24-NOV-1997) Plant Gene Expression Center, 800 Buchanan
                                      Street, Albany, CA 94710, USA
                         REFERENCE    3 (bases 1 to 59261)
                         AUTHORS      Theologis,A.
                         TITLE        Direct Submission
                         JOURNAL      Submitted (23-JAN-1998) Plant Gene Expression Center, 800 Buchanan
                                      Street, Albany, CA 94710, USA
                         REFERENCE    4 (bases 1 to 59261)
                         AUTHORS      Theologis,A.
                         TITLE        Direct Submission
                         JOURNAL      Submitted (23-JUN-1998) Plant Gene Expression Center, 800 Buchanan
                                      St., Albany, CA 94710, USA
                         COMMENT      On Jun 23, 1998 this sequence version replaced gi:2804593.
                                      The sequence of BAC T12M4 from Arabidopsis thaliana chromosome 1.
                         FEATURES             Location/Qualifiers
                         source          1..59261
                                         /organism="Arabidopsis thaliana"
                                         /mol_type="genomic DNA"
                                         /cultivar="Columbia"
                                         /db_xref="taxon:3702"
                                         /chromosome="1"
                                         /clone="T12M4"
                         gene            996..4040
                                         /gene="T12M4.1"
                         CDS             join(996..1187,1532..1693,2159..2494,
                                         2581..2662,2741..2823,3006..3045,3122..3221,3334..4040)
                                         /gene="T12M4.1"
                                         /note="Contains similarity to dihydrofolate reductase
                                         (dfr) gb|L13703 from Schizosaccharomyces pombe. ESTs
                                         gb|N37567 and gb|T43002 come from this gene."
                                         /codon_start=1
                                         /evidence=not_experimental
                                         /protein_id="AAC24078.1"
                                         /db_xref="GI:3249095"
                                         /translation="MGTSSCGDHEKQRIEDEEQYGVLLYYKYTSVPDLDELVSFVESS
                                         CNSLGLLGVRLSPKGVNVTVGGKLTALEBHIAAAKSNCLFEGTDFKLASCHHPLNDK
                                         VAEBCGFTSLSIRVVEKFQMLGSISLLQSFIRSCRVPLITFPCNSAEHLDGDGKSENK
```

```
                         /db_xref="GI:21954085"
                         /translation="IYGRDGYDFDGCRLRVEIAHGGRRFSPSVDRYSSSYSASARAPSR
                         RSDYHVLVTGLPPSASWQDLKDHMRKAGDVCFSEVFPDRKGMSGVVDYSNYDDMKYAI
                         RKIDATEFRNAFSSAYIRVREYESRSVSRSGPSDKSYSRSRSRGPSCSYSSKSRSVS
                         PARSISPRSRPLSRSRSLYSSVSRSQSRSKSRTRSRSNSPVSPVISG"
      misc_difference 1
                         /gene="At1g09140"
                         /note="compared to genomic sequence"
                         /replace="c"
      misc_difference 148
                         /gene="At1g09140"
                         /note="compared to genomic sequence"
                         /replace="g"
      misc_difference 579
                         /gene="At1g09140"
                         /note="compared to genomic sequence resulting in an amino
                         acid sequence difference"
                         /replace="t"
      3'UTR          627..933
                         /gene="At1g09140"
ORIGIN

Alignment Scores:
Pred. No.:
Score:                 2.09e-47        Length:          933
Percent Similarity:    1032.50         Matches:         205
Best Local Similarity: 94.50%          Conservative:    1
Query Match:           94.04%          Mismatches:      1
DB:                    71.16%          Indels:          11
                       8               Gaps:            1

US-10-014-927-19 (1-279) x AY050912 (1-933)

Qy    62  IleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHis  81
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3  ATTTATGGACGTGATGGTTATGATTTTGATGGGTGTCGACTTCGGGTTGAGATTGCACAT  62

Qy    82  GlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArg  101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63  GGTGGTCGTAGATTTCACCATCAGTTGATAGGTACAGCAGCAGCTACAGTGCGAGCCGT  122

Qy   102  AlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSer  121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123  GCACCCTCAAGACGCTCTGACTACCACGTGCTTGTGACGGGATTACCGGCCTTCTGCTTCG  182

Qy   122  TrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhe  141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183  TGGCAGGACCTTAAGGATCACATGCGCAAAGCTGGAGATGTCTGCTTCTCTGAAGTTTTC  242

Qy   142  ProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyr  161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243  CCTGACCGTAAAGGCATGTCTGGGGTTGTGGATTATAGCAACTATGATGATGAAGTAC  302

Qy   162  AlaIleArgLysIleAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArg  181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303  GCAATAAGGAAACTTGATGCCACTGAATTTCGAAATGCTTTCTCCAGTGCTTATATACGG  362

Qy   182  ValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerArgSerArgTyrArg  201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   363  GTGAGGGAATATGAGTCGAGGAGTGTGAGTCGAAGCCCAGATGATTCTAAAAGCTATAGA  422

Qy   202  SerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer  221
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   423  AGCAGGAGTCGGAGCCGTGCCGTGTCCAAGCTGTAGCTATAGTAGCAAGAGCGAGTGTGTCA  482

Qy   222  ProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSer  241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   483  CCTGCTAGATCCATTTCCCCGCGTTCACGGCCCCTTAGTCGTTCGTCGCTTACGTCGTATACAGC  542

Qy   242  SerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSer  261
          ||||||||||||||
Db   543  TCTGTCTCAAGG----------------------TCCCAATCAAGATCA  569

Qy   262  LysSerArgSerArgSerArgSerAsnSerProValSerProValIleSerGly  279
```

```
                    ELVLLDARNLYETRIGKFESENVETLDPEIRQYSDLPTWIDQNAEKMKGKNVLMYCTG
                    GIRCEMASAYIRSKGAGFENTFQLYGGIQRYLEQFPSSGGFFKGKNFVPDHRLCRMLVL
                    VCNHCRVKGDIYICELCRKHGKGKBVPLSLDPLNQPSESNGDNTRRKLRILCLHGFRQN
                    ASSFKGRTGSLAKKUKNIAELVFIDAPHELQFIYQTATPPSGGVCNKKPAWLVSSDFDK
                    PSETGWTVAQCQFDPLQYQTQTEGFDKSLTYLKTAFEKGPPDGILGFSQGAAMAAAV
                    CGKOEQLVGBIDRRFCVLCSGFTWPLLEMKEKRSIKCPSLHIFGSQPGKDRQIVTQA
                    SSDLAGLFEDGCATIVEHDFPGHIIPTKSPYIDEIKAFLYQFI"

gene                complement(4416. .7743)
                    /gene="T12M4.2"

CDS                 complement(join(4416. .4794,4920. .5047,5320. .5448,
                    5533. .5611,5859. .6022,6171. .6356,6581. .6712,6813. .6904,
                    7006. .7102,7513. .7743))
                    /gene="T12M4.2"
                    /note="Match to mRNA for importin alpha-like protein 4
                    (impa4) gb|Y14616 from A. thaliana. ESTs gb|N96440,
                    gb|N37503, gb|N37498 and gb|T42198 come from this gene."
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC24079.1"
                    /db_xref="GI:3249096"
                    /translation="MSLRPSTRAELRKKIYKTGVDADBARRRREDNLVEIRKNKREDS
                    LLKKRREGVMLQQQLPLGAGLDGFCTAAAVEKRLEGIPMMVQGVYSDDPQAQLEATTQ
                    FRKLLSIERSPPIDEVIKAGVIRPFVEFLGRHDHPQLQFEAAWALTNVASGTSDHTRV
                    VIEQGAVPIFVKLLTSASDDVREQAVWALGNVAGDSPNCRNLVLNYGALEPLLAQLNE
                    NSKLSMLRNATWTLSNFCRGKPPTPFEQVKPALPILRQLIYLNDEEVLTDACWALSYL
                    SDGPNDKIQAVIRAGVCPRLVELLGHQSPTVLIPALRTVGNIVTGDDSQTQRFIIESGV
                    LPHIYNLITQNHKKSSIKKEACWTTSNITAGNKLQIEAVVGAGIILPLVHLLQNAEFDI
                    KKEAWAISNATSGGSHEQIQYVLTCGCIKPLCDLLICPDPRIVTVCLEGLENILKVG
                    EADKEMGLNSGVNLYAQIIEESGLDKVENLQSHDNNEIYEKAVKILERYWAEEEEEQ"

gene                complement(8429. .8845)
                    /gene="T12M4.3"

CDS                 complement(8429. .8845)
                    /gene="T12M4.3"
                    /note="Contains similarity to DNAJ homologue gb|D84222
                    from Thermus thermophilus."
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC24080.1"
                    /db_xref="GI:3249097"
                    /translation="MNDSMKAPKFFKNQLKLABAHYQRGNLKGAIEIVNDLTFGHPNT
                    SSNHHEISQILLAYQINLTSQXASFTHYDILRISNPFCSHQMIQRKYRDILVKLYPDT
                    NKSIAAKSAFEINYAWKILSSDPEKRKDYNIKKRFK"

gene                complement(12117. .12740)
                    /gene="T12M4.4"

CDS                 complement(12117. .12740)
                    /gene="T12M4.4"
                    /note="ESTs gb|T04610, gb|N38459, gb|T45174, gb|R30481 and
                    gb|N64971 come from this gene."
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC24081.1"
                    /db_xref="GI:3249098"
                    /translation="MVESLFPSIENTGESSRRKPRISETABAEIBARRVNEESLKRW
                    KTNRVQQIYACKLVEALRRVRQRSSTTSNNETDKLVSGAARIRDTADRVLAASARGT
                    TRWSRAIIASRVKALLKHCHRKAKKSTGNCKSRKGLTETNRIKLPAVERKLKILGRLVP
                    GCRKVSVPNLLDEATDYIAALEMQVRAMEALLAELLTAAAPRTTLGT"

gene                complement(16337. .17299)
                    /gene="T12M4.5"

CDS                 complement(16337. .17299)
                    /gene="T12M4.5"
                    /note="EST gb|T21244 comes from this gene."
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC24082.1"
                    /db_xref="GI:3249099"
                    /translation="MGCCQDEQLVQTICDLYEKISKLESLKPSEDVNILFKQLVSTCIP
                    PNPNIDVTKMCDRVQEIRLNLLKICLGLAEGHLENHFSSILTSYQDNPLHHLNIPPYYN
                    NYLKLGKLBFDLLEQNLNGFVPKSVAFIGSGPLPLTSIVLASFHLKDTIFHNFDIDPS
                    ANSLASLLVSSDPDISQRMFFHFVDIMDVTESLKSFDVVLQGFEVLSIYHPTDDVINSVVISKKHP
                    LQKHMAPGAVLMLRSAHGPRAFLYPIVEPCDLQGFEVLSIYHPTDDVINSVVISKKHP
                    VVSIGNVGGPNSCLLKPCNCSKTHAKMNKNMMIEEFGAREEQLS"
```

```
CDS                 /gene="T12M4.6"
                    join(19686. .19873,20088. .20165,20262. .20471,20627. .20747,
                    20851. .20991,21237. .21398,21471. .21644,21951. .22038,
                    22169. .22227,22307. .22364,22526. .22650)
                    /gene="T12M4.6"
                    /codon_start=1
                    /evidence=not_experimental
                    /product="T12M4.6"
                    /protein_id="AAC24093.1"
                    /db_xref="GI:3249110"
                    /translation="MAAHATSEPAVNLPATQFESQVTEPFGVTLLVRHLPDGIPHDIV
                    SRLFSQYGASAVRPCSGGKLRNAAFVDFKNEAFASQAHRQLNGYQLRFPLGKVLQVQRA
                    NKPNDNKKSRQIEBSVTKGNAFSTVSNNDSKSGQIISGEPIAPKLGIDYPPPHLQY
                    AYPPPDANILANINALIAVPPLYTQVSFAVTVYLWIEVLHLMNKMNLPPPRLALP
                    TPPLPKAGPQQTDLEHQSSSESEMESDEDIGTSKSGKRARHGFLVGLGMDKDVPHET
                    VGVKPSSLTPKIPRIRKNHVMQIKITSKVTQDEYKEESENEDPADEPKEKDSNLKP
                    FASLELEKGRLPPQDILSLEMFKNYTAGNPSVVLYIKNLAKDVVIDDFYIFGSQFE
                    SSEVAKSSLGVRLMQEGRMRQQAFLITPPSVEVAHQSCKWFCVQRQTNDNPVWQDSWSS
                    QAKRMMMIYITSTVFEE"

gene                22919. .24297
                    /gene="T12M4.7"

CDS                 join(22919. .23125,23266. .24297)
                    /gene="T12M4.7"
                    /codon_start=1
                    /evidence=not_experimental
                    /product="T12M4.7"
                    /protein_id="AAC24094.1"
                    /db_xref="GI:3249111"
                    /translation="MNQTLKSSTSSIHTSPLLGFFSFIRSKTLGNFFYSILCFDAILL
                    AKPLFMLTSSTNSNDSISCLITIRTALVGMYLVGGMIDAHKVFDEMPERNPVTWNV
                    MITGLTNLGDFBKALCPLEKMPNRTVSWTTIIDGYARVDKPKEALLSRMVACDAI
                    KPNEITTIAILPAVWNLGDLKMCGSVHAYVGKRGFVPCDIRVTNSLIDAYAKCGCIQS
                    AFKFFIEIPNGRKNLVSWTTMISAFAIHGMCKEAVSMFKDMERLGKPNRVTMISVLN
                    ACSHGGLAEEEFLEFFNTMVNEYKITPDVKHYGCLVDMLRRKGRLEAEKIALEIPIE
                    EKAVVWRMLLGACSVYDDAELABRVTRKLMELERSHGGDYVLMSNIFCGTGRFLDAQR
                    FRKQMDVRGVAKLPGHSQVT"

gene                25594. .28968
                    /gene="T12M4.8"

CDS                 join(25594. .25687,26293. .26400,26495. .26687,26785. .27043,
                    27217. .27321,27408. .27494,27577. .27634,28142. .28209,
                    28295. .28390,28657. .28531,28693. .28722,28807. .28968)
                    /gene="T12M4.8"
                    /note="Match to calreticulin (AtCRTL) mRNA gb|U27698 and
                    DNA gb|U66344. ESTs gb|T45719, gb|T22451, gb|H36323 and
                    gb|AA042519 come from this gene."
                    /codon_start=1
                    /evidence=not_experimental
                    /protein_id="AAC24083.1"
                    /db_xref="GI:3249100"
                    /translation="MAKMIPSLVSLILIGLVAIASAAVIFEERFPDGWENRWVKSEWK
                    KDDNTAGEWKHTAGNWSGDANDKGIQTSEDYRFYAISAEFPEFSNKDKTLVFQFSVTGH
                    EQKLDCGGGYMKLLSGDVDQKKFPGGDTPYSIMFGPDICGYSTKKVHAILTYNEANHLI
                    KKDVPCETDQLTHVTFILRPDATYSILLIDNVEKQGTGSLYSDWDLLPFKKIHDSAKK
                    PEDWDEQEYISDPEDKKPDGYDDIPKEIPDTDSKKPEDWDDEEDGEWTAPTIPNPEYM
                    GEWKPQIKNPNYKGKWEAPLIDNPDFKDDPDLYVFPKLKYVGLELWQVKSGSLFDNV
                    LICDDPYAKLADETWGKLKDNIOSNITLQAEKRAAFDAEREEESKDAPAESDA
                    EDEPEDDEGGDDSDESKAEETKSVDSETSEKDATAHVCVHNITWRNFTHP"

gene                30117. .30527
                    /gene="T12M4.9"
```

```
Alignment Scores:
Pred. No.:             59261
Score:                 904.00          Length:           59261
Percent Similarity:    43.68%          Matches:          242
Best Local Similarity: 43.68%          Conservative:     0
Query Match:           62.30%          Mismatches:       3
DB:                                    Indels:           312
                                       Gaps:             9

                       1.86e-38

US-10-014-927-19 (1-279) x T12M4 (1-59261)

Qy    1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgIleArgLys 20

Db 56431 ATGAGTAGCCGATGGAATCGCACGATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG 56490
        19686. .22650
```

```
Qy     21 CysGluValGluAspLeuPheTyrLys---------                                          29
Db  56491 TGTGAAGGTTGAAGATCTCTTCTACAAGGTTGAAAATTCCTCTTTCTCGATAAAAT                      56550

Qy     29 ---------                                                                     29
Db  56551 TGAATTCATTATGACTTGTTGGGTTCATAAATTGCAATTCTGCTTGCTGCGAGACAATT                   56610

Qy     30 ---------TyrGlyProIleValAspIleAspLeuLys                                       39
Db  56611 TAAATGACTCTTATGTATATTTGTTCAGTAATGGGACCAATTGTGGACATTGATTGAAG                   56670

Qy     40 IleProProArgProProGlyTyrAlaPheValGlu---------                                 51
Db  56671 ATTCCACCGAGACCCTCCTGGTTATGCCTTGTCGA-GTATATTGATCAAGTACAAATTT                   56729

Qy     51 ---------                                                                     51
Db  56730 GTTTTTTCTTCTCTGTAAATGTAATAGGCTAATGACTAAGATAGTTGTTATTGGTGG                     56789

Qy     52 ---PheGluAspProArgAspAlaAspAlaAlaIleTyrGlyValArgAspGlyTyrAspPhe               70
Db  56790 CAGTTGAAGATCCTCGTGATGCAGACGATGCAATTTATGGACGTGATGGTTATGATTTT                   56849

Qy     71 AspGlyCysArgLeuArg---------                                                   76
Db  56850 GATGGGTGTCGACTTCGGGTTAGTAGTAAACGCATGATGAAAGCTAGCTTAATTTCTGTAAT               56909

Qy     77 ---------ValGluIleAlaHisGlyGlyAr                                              84
Db  56910 TTCTTGTAAAGGTGTTATCTTTGTGATGTTTTTAGGTTGAGATTGCACCATGGTGGTCG                   56969

Qy     84 gArgPheSerProSerValAspArgTyrSerSerSerTyrSerSerArgAlaProSe                     104
Db  56970 TAGATTTTCACCATCAGTTGATAGGTACAGCAGCTACAGTGCGAGCCGTGCACCTTC                     57029

Qy    104 rArgArgSerArgTyrArg---------                                                  110
Db  57030 AAGACGCCTCGACTACCGGCGGGTTGTGAGCTGTTGTATTGGTGTTGTGTT                           57089

Qy    111 ---------ValLeuValThrGlyL                                                     116
Db  57090 AAAATTTTATATTTGAAAACTCATTTTTACTACCTAAACATGTAGTGCTTGTGACCGGAT                  57149

Qy    116 euProProSerAlaSerTrpGlnAspLeuLys---------                                     126
Db  57150 TACCGCCTTCTGCTTCGTGCGGACGACCTTAA-GGTAAGGGACACTATATAGTCTTTTCT                  57208

Qy    126 ---------                                                                     126
Db  57209 CTGAATGTTGGTTGTCTCTATATCATGTTTTGGATTTATCTCTTTTCTGAAATGATGTTATT               57268

Qy    127 ---------AspHisMetArgLysAlaGlyLysAspValCysPheSerGluV                          140
Db  57269 TGCTATTTACGGGTGATTAGGAATGCACACATGCCAAAGCTGGAGATGTCTGCTTCTCTGAAG              57328

Qy    140 alPheProAspArgLys---------                                                    145
Db  57329 TTTTCCCTGACCGTAAAGGTGAGTTGACATTCGATAGTTTGATTGATAAGCTTTGATTGATG               57388

Qy    145 ---------                                                                     145
Db  57389 TGTTAGTAAATTAGTCTTTGTGAGGAGAATAGGTTGTTAAGCATCGAACTGCTAAACTC                  57448

Qy    146 ---------GlyMetSerGlyValValAspTyrSerAsnTyrAspAspM                             159
Db  57449 ACATTCAGTATTCTTTGTAGGCATGTCCGGGGGTGTCGATTATAGCAACTATGATGATA                   57508

Qy    159 etLysTyrAla-                                                                  162
Db  57509 TGAAGTACGC-AGTAAGTTTTATATCTTTGCAACGCAAATGTTCCTGGACTTATGCCTTA                  57567


Qy    162 ---------                                                                     162
Db  57568 GACTGCTTTGTTCATAGTATACCGAGCTGAATTATCTTCCTGGAGGCCAGTGTTGGA                     57627

Qy    163 ---------IleArgLysLeuAspAlaThrGluP                                            171
Db  57628 TCTTTGATGTCCCTTAAAATTTTGATGTGCACAGATAAGGAAACTTGATGCCACTGAAT                   57687

Qy    171 heArgAsnAlaPheSerAlaTyrIleArg                                                 181
Db  57688 TTCGAAATGCTTTCTCTAGTGCTTATATACGGGTATGTTGATTGCTTTCTTTGATTTG                    57747

Qy    182 ---------ValArgGluTyrGlu                                                      186
Db  57748 TTAAGCATAAGTGGATAATGGAGTCATCTCTGAATTTACTGCTTCAGGTGAGGGAATATGAG               57807

Qy    187 SerArgSerValSerArgSerProAspAspSerLysSerArgSerArgSer                           206
Db  57808 TCGAGGAGTGTGAGTCGAAGCCCAGATGAATCTAAAAGCTATAGAGAGCAGGAGTCGGAGC                57867

Qy    207 ArgGlyProSerCysSerCysSerLysSerArg                                             218
Db  57868 CGTGGTCCAAGCTGTAGCTATAGTAGCAAGAGCAGGAGGTCTTTTTTTTTTTTTTTTTTT                  57927

Qy    218 ---------                                                                     218
Db  57928 CATAAACCTAAGACATATAAGGAGATTTTATTGTAACTTATTATTATTATGAAATTAACTGACTT           57987

Qy    219 ---------SerValSerProAlaArgSerIleSerProArgSerArgProLeuSe                      234
Db  57988 CTAAATGCAATGCAGTTGTCACCTGCTAGATCCATTTCCCGCGGTTCACGGCCCCTTAG                   58047

Qy    234 rArgSerArgSerLeuTyrSerSerValSerArg                                            245
Db  58048 TCGTTCTTCGCTTCGCCTATACAGCTCTGTCTCAAGG                                         58081
```

RESULT 5
ATH131214
LOCUS   ATH131214   5164 bp   DNA   linear   PLN 07-MAY-1999
DEFINITION Arabidopsis thaliana srp30 gene, exons 1-12.
ACCESSION   AJ131214
VERSION   AJ131214.1 GI:4775269
KEYWORDS   SF2/ASF-like splicing modulator; srp30 gene.
SOURCE   Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Lopato,S., Kalyna,M., Dorner,S., Kobayashi,R., Krainer,A.R. and
            Barta,A.
  TITLE   atSRp30, one of two SF2/ASF-like proteins from Arabidopsis
            thaliana, regulates splicing of specific plant genes
  JOURNAL   Genes Dev. 13 (8), 987-1001 (1999)
  MEDLINE   99234087
   PUBMED   10215626
REFERENCE   2 (bases 1 to 5164)
  AUTHORS   Barta,A.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-DEC-1998) Barta A., Institute for Biochemistry,
            University of Vienna, Dr. Bohrgasse 9/3., A-1030 Vienna, AUSTRIA
FEATURES   Location/Qualifiers
     source   1..5164
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
     mRNA   join(1806..2014,2138..2203,2290..2364,2444..2546,
                     2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
                     3501..3578,4515..4572,4844..5125)
     exon   1806..2014
                     /number=1
     gene   1928..4858
                     /gene="srp30"

```
CDS             join(1928..2014,2138..2203,2290..2364,2444..2546,
                2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
                3501..3578,4515..4572,4844..4858)
                /gene="srp30"
                /codon_start=1
                /product="SF2/ASF-like splicing modulator Srp30"
                /protein_id="CAB42557.1"
                /db_xref="GI:4775270"
                /db_xref="GOA:Q9XFR5"
                /db_xref="SPTREMBL:Q9XFR5"
                /translation="MSSRWNRTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPRP
                PGYAFVEFEDPRDADDAIYGRDGYDFDGCRLRVEIAHGGRRFSPSVDRYSSSYSASRA
                PSRRSDYRVLVTGLPPSASWQDLKDHMRKAGDVCFSEVFPDRKGMSGVVDYSNYDDMK
                YAIRKLDATEFRNAFSSAYIRVREYESRSVSRGPDDSKSYRSRSRGPSCSYSSKSR
                SVSPARSISPRSRRPLSRSRSLYSSVSRSQSRSKSRSNSPVSPVISG"
CDS             join(1928..2014,2138..2203,2290..2364,2444..2546,
                2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
                3501..3578,4178..4214)
                /gene="srp30"
                /note="alternative exon 10a"
                /codon_start=1
                /product="SF2/ASF-like splicing modulator Srp30, variant
                1"
                /protein_id="CAB42558.1"
                /db_xref="GI:4775271"
                /db_xref="GOA:Q9XFR6"
                /db_xref="SPTREMBL:Q9XFR6"
                /translation="MSSRWNRTIYVGNLPGDIRKCEVEDLFYKYGPIVDIDLKIPPRP
                PGYAFVEFEDPRDADDAIYGRDGYDFDGCRLRVEIAHGGRRFSPSVDRYSSSYSASRA
                PSRRSDYRVLVTGLPPSASWQDLKDHMRKAGDVCFSEVFPDRKGMSGVVDYSNYDDMK
                YAIRKLDATEFRNAFSSAYIRVREYESRSVSRGPDDSKSYRSRSRGDVCSYSSSKSR
                SVSPARSISPRSRRPLSRSRSLYSSVSRSGSLLRAGDWI"
intron          2015..2137
                /gene="srp30"
exon            2138..2203
                /gene="srp30"
                /number=1
intron          2365..2443
                /gene="srp30"
                /number=2
exon            2204..2289
                /gene="srp30"
                /number=2
intron          2547..2631
                /gene="srp30"
                /number=4
exon            2290..2364
                /gene="srp30"
                /number=3
exon            2632..2678
                /gene="srp30"
                /number=5
intron          2679..2785
                /gene="srp30"
                /number=5
exon            2444..2546
                /gene="srp30"
                /number=3
exon            2786..2843
                /gene="srp30"
                /number=6
intron          2844..2967
                /gene="srp30"
                /number=6
exon            2968..3017
                /gene="srp30"
                /number=7
intron          3018..3160
                /gene="srp30"
                /number=7
exon            3161..3217
                /gene="srp30"
```

```
intron          3218..3290
                /number=8
                /gene="srp30"
intron          3291..3403
                /number=8
                /gene="srp30"
                /number=9
intron          3404..3500
                /gene="srp30"
                /number=9
exon            3501..3578
                /gene="srp30"
                /number=10
intron          3579..4514
                /gene="srp30"
                /number=10
exon            4178..4572
                /gene="srp30"
                /note="alternative exon"
                /number=11
exon            4178..4340
                /gene="srp30"
                /note="alternative exon 10a"
                /number=11
exon            4515..4572
                /gene="srp30"
                /number=11
intron          4573..4843
                /gene="srp30"
                /number=11
exon            4844..5125
                /gene="srp30"
                /number=12
```

```
ORIGIN

Alignment Scores:
Pred. No.:                                    Length:                5164
Score:                     5.92e-39           Matches:                242
Percent Similarity:        891.00             Conservative:             0
Best Local Similarity:     43.60%             Mismatches:               3
Query Match:               43.60%             Indels:                 313
DB:                        61.41%             Gaps:                     9
                           8

US-10-014-927-19 (1-279) x ATH131214 (1-5164)

Qy      1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1928 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG 1987

Qy     21 CysGluValGluAspLeuPheTyrLys                                   29
          |||||||||||||||||||||||||||
Db   1988 TGTGAGGTTGAAGATCTCTTCTACAAGGTTGAAAATTCCTCTTTCTCTGATAAAAAT   2047

Qy     29                                                              29
Db   2048 TGAATTCATTATGACTAGTTTGGGTTCATAAATTCGCAATTCGTCTTGCTGAGACAATT 2107

Qy     30                                           TyrGlyProIleValAspLeuLys 39
                                                    ||||||||||||||||||||||||
Db   2108 TAAATCGACTCTTATGTATTGTTTCAGTATGGACCAATTGTGGACATTGATTTGGAAG  2167

Qy     40 IleProProArgProProGlyTyrAlaPheValGlu                          51
          |||||||||||||||||||||||||||||||||||
Db   2168 ATTCCACCGAGACCCTCCTGGTTATGCCTTGTCGA-GGTATATTGATCAAGTACAAATTT 2226

Qy     51                                                              51
Db   2227 GTTTTTTCTTCTTCTTGTAATAGTATAGGCTAATGACTAAGATAGTTGTTATTGGTGG  2286

Qy     52 --FheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPhe      70
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2287 CAGTTGAAGATCCTCGTGATGCAGACCAATTATGGACGTGATGGTTATGATTTT      2346

Qy     71 AspGlyCysArgLeuArg                                            76
          |||||||||||||
```

Mon Mar 29 09:55:42 2004   us-10-014-927-19.rge   Page 8

```
Qy    219  ------------------------SerValSerProAlaArgSerIleSerProArgSerArgProLeuS  234
Db   3485  TCTAAATGCAATGCAGTGTGTCACCTGCTAGATCCATTCCCGCGTTCACGGCTCCCCTTA           3544

Qy    234  erArgSerArgSerLeuTyrSerSerValSerArg  245
Db   3545  GTCGTTCTCGCTCGCTATACAGCTCTGTCTCAAGG  3579


RESULT 6
AX040661
LOCUS        AX040661                  4044 bp     DNA        linear    PAT  23-NOV-2000
DEFINITION   Sequence 18 from Patent WO0065059.
ACCESSION    AX040661
VERSION      AX040661.1  GI:11340389
KEYWORDS     .
SOURCE       unidentified
  ORGANISM   unidentified
             unclassified.
REFERENCE    1
  AUTHORS    Barta,A., Lopato,S., Kalyna,M. and Dorner,S.
  TITLE      Splicing factor
  JOURNAL    Patent: WO 0065059-A 18 02-NOV-2000;
             \sterreichisches Forschungszentrum Seibersdorf Ges.m.b.H.;  (AT)
FEATURES             Location/Qualifiers
     source          1..4044
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
                     /note="Beschreibung des unbekannten ORGANISM:Genom
                     atSRp30"

ORIGIN
Alignment Scores:
Pred. No.:                3.26e-38      Length:            4044
Score:                    875.00        Matches:           242
Percent Similarity:       43.53%        Conservative:      0
Best Local Similarity:    43.53%        Mismatches:        3
Query Match:              60.30%        Indels:            314
DB:                                     Gaps:              9

US-10-014-927-19 (1-279) x AX040661 (1-4044)

Qy      1  MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys       20
Db    808  ATGAGTAGCCGATGGAATCGCACCATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG      867

Qy     21  CysGluValGluAspLeuPheTyrLys                                       29
Db    868  TGTGAGGTTGAAGATCTCTTCTACAAGGTTTGAAAATTCCTCTTTTCTCGATAAAAA        927

Qy     29                                                                    29
Db    928  TTGAATTCATTATGACTAGTTTGGGTGGTTCATAAAATTGCATTCGTCTTGCTGAGACAAT    987

Qy     30                                    TyrGlyProIleValAspIleAspLeuY       39
Db    988  TTAAATGCGACTCTTATGTATATTGTTTCAGTAGTATATTGGACATTGTGGACATTGATTGAA  1047

Qy     39  sIleProProGlyTyrAlaPheValGlu-                                    51
Db   1048  GATTCCACCGGAGACCTCCTGGTTTATGCCTTTGTCGA-GGTATATTGATCAAGTACAAATT  1106

Qy     51                                                                    51
Db   1107  TGTTTTTTTCTTCTTCCTTGTAATAGTATAGGCTAATGACTAAGAGTTGTTATTATTGGTG  1166

Qy     52  ----PheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPh        70
Db   1167  GCAGTTTGAAGATCCTCGTGATGCAGACGATGCAATTTATGGACGTGATGGTTATGCATTT  1226

Qy     70  eAspGlyCysArgLeuArg                                              76
Db   3425
```

```
Db   2347  GATGGGTGTCGACTTCGGGTTAGTAAACGATGATGAAAGCTAGCTTAATTTCTGTAAT      2406
Qy     77  ------------------------ValGluIleAlaHisGlyGlyArAr-              84
Db   2407  TTCTTGTAAAGGTGTATCTTGTGTGATGTTTTAGGTTGAGATTGCACATGGTGGTCG        2466

Qy     84  gArgPheSerProSerValAspArgTyrSerSerTyrSerArgAlaSerArgAlaProSe   104
Db   2467  TAGAATTTTCACCATCAGTTGATAGGTACAGCAGCTACAGTGCCGAGCCGTGCACCTTC      2526

Qy    104  rArgArgSerAspTyrArg------------------------------------------   110
Db   2527  AAGACGCCTGACTACCGCGGTTGTGTAGAGTCTTCTGATTGTGTTATTGGTGTGTGT        2586

Qy    111  ----------------------------------ValLeuValThrGlyL               116
Db   2587  AAAATTTTATATTGAAAACTCATTTTACTACCTAAACATGTAGTGCTTGTGACCGGAT       2646

Qy    116  euProProSerAlaSerTrpGlnAspLeuLys------------------------------  126
Db   2647  TACCGGCCTTCTGCTTCGTGTGGCCAGGACCTTAA-GGTAAGGGACACTATATAGTCTTTTTCT 2705

Qy    126  ------------------------------------------------------------   126
Db   2706  CTGAATGTTGGTTCTCTATATCATGTTTGGATTTATTCTTTTTCTGAAATGATGTTATT      2765

Qy    127  ------------AspHisMetArgLysAlaGlyValCysPheSerGluV               140
Db   2766  TGCTATTTACGGTGATTAGGATCACATGCGCAAAGCTGGACATGTCTGCTTCTCTGAAG      2825

Qy    140  alPheProAspArg-Lys------------------------------------------   145
Db   2826  TTTTCCCTGACCGTAAAGGTGAGTTGACATTCGATAGTTTGGATAAGCTTTTGATTGAT      2885

Qy    145  ------------------------------------------------------------   145
Db   2886  GTGTTAGTAAAATTAGTCTTTGTGAAGGAGAATAGGTGTTAAGCATCTGAACTGCTAAACT    2945

Qy    146  ------------GlyMetSerGlyValValAspTyrSerAsnTyrArgAspAsp           158
Db   2946  CACATTCAGTATTCTTTGTGTAGCCATGTCTGGGGGTGTGGATTTATAGCAACTATGATGAT   3005

Qy    159  MetLysTyrAla------------------------------------------------   162
Db   3006  ATGAAGTACGGC-AGTAAGTTTATATCTTTGCAACGCAAAGTTCCTGGACTTATGCCTT      3064

Qy    162  ------------------------------------------------------------   162
Db   3065  AGACTGCTTTGTTTCATAGTATATACCGAGCTGAATTATCTTCCTGGAGGCCAGTGTTGG     3124

Qy    163  ------------IleArgLysLeuAspAlaThrGlu                             170
Db   3125  ATCTTTGATGTTCCCTAAAATTTTGATGTGACAGATAAGGAAACTTGATCCCACTGAA       3184

Qy    171  PheArgAsnAlaPheSerSerAlaTyrIleArg---------------------------   181
Db   3185  TTTCGAAATGCTTTCTCCAGTGCTTATATACGGGTATGTTGTATTGCTTTCTTTGATTTT     3244

Qy    182  --------------------------ValArgGluTyrGl                        186
Db   3245  GTTAAGCATAAGTGAGTCAGATATGGAGTCATCCTGAATTTACTGTCTGTTCAGGTGAGGGAATGA 3304

Qy    186  uSerArgGlyValSerArgSerProAspAspSerLysSerArgSerArgSe            206
Db   3305  GTCCGAGGAGTGTGAGTCGGAGCCCAGATGATTCTAAAAGCTTATAAAGCTATAGAGCAGGAGTCGGAG 3364

Qy    206  rArgGlyProSerCysSerTyrArgSerLysSerArg------------------------  218
Db   3365  CCGTTGTCCAAGCTGTAGCTATAGTAGCAGAAGACGAGGAGGTCTTTTTTTTTTTTTTTT     3424

Qy    218  ------------------------------------------------------------   218
Db   3425  TCATAAACCTAAGACATATAAGGGATTTTTATTGTAACTTATTTATGAAATTAACTGACT     3484
```

```
Qy    219  --------------------------SerValSerProAlaArgSerIleSerProArgArgProLeu 233
Db   2365  TTCTAAATGCAATGCAGTGTGTCACCTGCTCAGATCCATTTCCCCGCGTTCACGGCCCCTT 2424

Qy    234  SerArgSerArgSerLeuTyrSerSerValSerArg 245
Db   2425  AGTGCGTTCTCGCTCGCTATACAGCTCTGTCTCAAGG 2460
```

RESULT 7
```
LOCUS       AK118379               1229 bp    mRNA          linear   PLN 06-DEC-2002
DEFINITION  Arabidopsis thaliana At1g02840 mRNA for putative ribonucleoprotein
            SF-2, complete cds, clone: RAFL19-64-H16.
ACCESSION   AK118379
VERSION     AK118379.1  GI:26451793
KEYWORDS    FLI cDNA; CAP trapper.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
            Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
            Kawai,J., Hayashizaki,Y. and Shinozaki,K.
  TITLE     Arabidopsis thaliana full-length cDNA
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 1229)
  AUTHORS   Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
            Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
            Kawai,J., Hayashizaki,Y. and Shinozaki,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
            Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
            230-0045, Japan (E-mail:mseki@gsc.riken.go.jp,
            URL:http://pfgweb.gsc.riken.go.jp/, Tel:81-45-503-9625,
            Fax:81-45-503-9586)
COMMENT     An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al. (1998) Plant J. 15:707-720;
            Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1-B vector (Carninci et
            al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
            This clone is in a modified pBluescript vector.
            Please visit our web site (http://pfgweb.gsc.riken.go.jp/) for
            further details.
FEATURES             Location/Qualifiers
     source          1..1229
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="RAFL19-64-H16"
                     /note="common name: thale cress"
     gene            1..1229
                     /gene="At1g02840/F22D16_30"
     CDS             93..1004
                     /gene="At1g02840/F22D16_30"
                     /codon_start=1
                     /product="putative ribonucleoprotein SF-2"
                     /protein_id="BAC42991.1"
                     /db_xref="GI:26451794"
                     /translation="MSSRSSRTVYVGNLPGDIREREVEDLFSKYGPVVQIDLKVPPRP
                     PGYAFVEFDDARDAEDAIHGRDGYDFDGHRLRVELAHGGRRSSDDTRGSFNGGRGGGG
                     RGRGDGSGRGPSRRSRERFVLVTGLPSSASWQDLKDHMRKGGDVCFSQVYRDARGTTGV
                     VDYTCYEDMKYALKKLDDTEFRNAFSNGYVRVREYDSRKDSRSPSRGRSYSRKSRSRSR
                     GRSVSRSRSRSRSRSRSPSKAKSSRSPSRAKSTRSSPGPRSKRSRSPRRSRSRSRSPLP
                     SVQKEGSKSPSKPSPAKSPIHTRSPSR"
ORIGIN
Alignment Scores:
Pred. No.:               2.93e-38          Length:        1229
```

```
Db   1227  TGATGGGTGTCGGGTTAGTAAACGCATGATGAAAGCTAGCTTAATTTCTGTAA 1286

Qy     77  ---------------------------ValGluIleAlaHisGlyGlyAla 84
Db   1287  TTTCTTGTGTAAAGGTGTTATCTTCGTGTGATGTTTTTAGGTTGAGATTGCACATGTGTGGTC 1346

Qy     84  rgArgPheSerProSerValAspArgTyrSerSerAlaSerArgAlaProS 104
Db   1347  GTAGATTTTCACCATCAGTTGATAGGTACAGCAGCAGTACCAGTGCGAGCCGTGCACCTT 1406

Qy    104  erArgArgAspTyrArg--------------------------------- 110
Db   1407  CAAGAGCGCCTCGACTACCACGGCCGGTTTGTAGAGCTTCTCGATTGTGTATTGGTGTGTG 1466

Qy    111  -----------------------------------ValLeuValThrGly 115
Db   1467  TAAAAATTTTATATTTGAAAAACTCATTTTTACTACCTAAACATGTAGTGCTTGTGACCGGA 1526

Qy    116  LeuProProSerAlaSerTrpGlnAspLeuLys----------------- 126
Db   1527  TTACCGCCTTCTGCTTCGTGGCAGGACCCTTAA-GGTAAGGGACACTATATAGTCTTTTTC 1585

Qy    126  ------------------------------------------------- 126
Db   1586  TCTGAATGTTGGTTCTCTATATCATGTTGGATTTATCTCTTTCTGAAATGATGTTAT 1645

Qy    127  ----------AspHisMetArgLysValAlaGlyAspValCysPheSerGlu 139
Db   1646  TTGCTATTTACGGGTGATTAGGATCACATGCCAAAGCTCGAGATGTCTCGTCTTCTCTGAA 1705

Qy    140  ValPheProAspArg-Lys-------------------------------- 145
Db   1706  GTTTTTCCCTGACCGTAAAGGTGAGTTGACATTCGATAGTTTGGATAAGCTTTTTGATTGA 1765

Qy    145  ------------------------------------------------- 145
Db   1766  TGTGTTAGTAAATTAGTGCTTTGTGAGGAGAATAGGTGTTAAGCATCTGAACTGCTAAAC 1825

Qy    146  ----------GlyMetSerGlyValValAspTyrSerAsnTyrAspAsp 158
Db   1826  TCACATTCAGTATTCTTTGTAGGCATGTCTGGGGTGTGATTATAGCAACTATGATGA 1885

Qy    158  pMetLysTyrAla------------------------------------- 162
Db   1886  TATGAAGTACGC-AGTAAGTTTTATATCTTGCAACGCAAATGTTCCTGACTTATGCCT 1944

Qy    162  ------------------------------------------------- 162
Db   1945  TAGACTGCTTTGTTCATAGTATATACCGAGCTGAATTTATCTTCCTGGAGGCCAGTGTTG 2004

Qy    163  ----------------IleArgLysLeuAspAlaThrGl 170
Db   2005  GATCTTTGATGTTCCCTAAAATTTTTGATGTGACAGATAGGAAACTTGATGTCCCACTGA 2064

Qy    170  uPheArgAsnAlaPheSerSerAlaTyrIleArg--------------- 181
Db   2065  ATTTTCGAAATGCTTTCTCTAGTGCTTATATACGGGTATGTTGTATTGCTTTCTTGATTT 2124

Qy    182  -----------------------------ValArgGluTyrArg 186
Db   2125  TGTTAAGCATAAGTGGATAATGGAGTCATCTCGAATTTACTGTTCAGGTGAGGAATATG 2184

Qy    186  luSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgS 206
Db   2185  AGTCGAGGAGTGGAGTCCGAGCCCAGATGAATTCTAAAAGCTATAGAAGCAGGAGTGCGGA 2244

Qy    206  erArgArgGlyProSerCysSerTyrSerGluLysSerArg--------- 218
Db   2245  GCCGTGGTCCAAGCTGTCCAGTCTATAGTAGCAAGCACGAGCGAGGTCTTTTTTTTTTT 2304

Qy    218  ------------------------------------------------- 218
Db   2305  TTCATAAACCTAAGACATATAAGGGATTTTTTATTGTAACTTATTTATTGAAATTAACTGAC 2364
```

```
Score:                  865.00      Matches:        184
Percent Similarity:     74.32%      Conservative:    33
Best Local Similarity:  63.01%      Mismatches:      45
Query Match:            59.61%      Indels:          30
DB:                                 Gaps:             8
```

US-10-014-927-19 (1-279) x AK118379 (1-1229)

```
Qy    1   MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys  20
              |||||||||||  ||||       ::::|||
Db   93   ATGAGCAGTCGTTCGAGTAGAACCGTGTACGTCGGAAACCTTCCTGGCGATATCCGTGAG  152

Qy   21   CysGluValGluAspLeuPheTyrLysTyrGlyTyrProIleValAspIleAspLeuLysIle  40
              ||||   |||||||       :::|||       ::|||
Db  153   AGAGAGGTCGAAGATTTGTTCAGTAAGTATGGATATCCGATTGTTGATATTGATTTGAAGGTT  212

Qy   41   ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp  60
              ||||||||||||||       ||:||       :::|||
Db  213   CCTCCAAGGCCTCCTGGTTATGCTTTCGTTGAGTTTGATGATCCTCGGGATGCTGAAGAT  272

Qy   61   AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleGluAla  80
              |||       |||||||       :::|||       ||:|||
Db  273   GCTATTCATGGTCGTGATGGCTATGACTTTGATGGGCATCGTTTGAGGGTGGAATTGGCG  332

Qy   81   HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSer------  98
              ||||:::       :::|||
Db  333   CATGGTGGGGAGGCGT--------TCATCAGATGATACTCGGGGTAGTTCAATGGTGGT  383

Qy   99   ------------------------AlaSerArgAlaProSerArgArgSer  107
                                          :::|||
Db  384   GGCCGTGGTGGTGGTGTCGCGGCCCTGGTGATGGGTAGTCGTGGCCATCGTAGGAGATCA  443

Qy  108   AspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAsp  127
              ::::|||       |||||||       ||:|||
Db  444   GAGTTTCGTGTTCTAGTCACAGGTTGCCTTCATCGCTTCTTGGCAAGATCTCAAGGAT  503

Qy  128   HisMetArgArgAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMet  147
              ||||:::       :::|||       ||:|||
Db  504   CACATGCGAAGAGCGGCGATGTCTGTTTCTGCAAGTGTACCGTGATGCCTAGAGGACCA  563

Qy  148   SerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleLysArgLysLeuAsp  167
              ||||:::       :::|||       |||:|||
Db  564   ACTGGAGTTGTTGATTACACCTGCTATGAGGACATGAAGAAGTAGCGCTGAAAAAGCTGAC  623

Qy  168   AlaThrGluPheArgAsnAlaPheSerArgAlaTyrIleArgValArgGluTyrGluSer  187
              ||||:::       |||||||       :::|||
Db  624   GCACCAGAGTTTCGAAATGCGTTTTCGAATGGATATGTCCGGGTTAGAGAATATGATTCA  683

Qy  188   ArgSerValSerArgSerProAspAspSerLysSerSerTyr---ArgSerArgSer  206
              ||||:::
Db  684   AGGAGGGATTTCTAGGAGTCCTAGCCGGGAAGATCCTATTCTAAGAGCCGCAGCCGCAGC  743

Qy  207   ArgGlyProSerCysSerArgSerLysSerArgArgSerArgSerValSerProAlaArgSerIle  226
              ||||:::       :::|||
Db  744   CGTGGACGAAGCGTG--------AGCCGAAGCAGGAGCAGGAGCAGGAGGAGCAGA  794

Qy  227   SerProArgSerArgProLeuSerArgArgSerArgSerLeuTyrSerVerSerArgSer  246
              |||:::::::
Db  795   AGTCCAAGGCAAAG------TCTTCACGTAGGTCCCCTGCAAAATCTCACATCCAAGATCT  848

Qy  247   GlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSer------Arg  264
              ||||
Db  849   CCTGGCCCCCGC---------------TCGAAGTCAAGGTCACCGTCTCCAAGAAGA  890

Qy  265   SerArgSerArgSerAsnSerProValSerProVal  276
              ||||:::
Db  891   TCGCGTTCAAGATCAAGATCTCCTCTACCTTCTGTT  926
```

REFERENCE AUTHORS and other GenBank fields:

```
VERSION     AY085920.1  GI:21404630
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1232)
AUTHORS     Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
            Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
TITLE       Full-length messenger RNA sequences greatly improve genome
            annotation
JOURNAL     Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE     22088475
PUBMED      12093376
REFERENCE   2  (bases 1 to 1232)
AUTHORS     Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
TITLE       Full-Length cDNA from Arabidopsis thaliana
JOURNAL     Unpublished
REFERENCE   3  (bases 1 to 1232)
AUTHORS     Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
TITLE       Direct Submission
JOURNAL     Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
            Malibu, CA 90265, USA
COMMENT     This clone sequence is one of 5,000 Ceres full-length cDNAs made
            available to TIGR and Genbank. The following quality assessment of
            this set was done by comparison with known proteins: two percent of
            the clones are estimated to be 5'-truncated; less than one percent
            are 3'-truncated; approximately two percent represent alternative
            splice variants, including unspliced introns and spliced exons; one
            percent may contain premature stop codons; five percent may have
            frame shifts in a coding region. A sequence is considered to be
            5'-truncated if it lacks the translation initiation start (ATG). A
            sequence is considered to be 3'-truncated if it lacks the
            C-terminal end of the encoded protein. Please note that these cDNA
            sequences are derived from the Ws or LAer ecotypes and therefore
            may contain polymorphisms when compared to sequences from Col-O.
            Genset carried out the library production and sequencing of the
            full-length clones. Ceres, Inc. carried out the clustering of the
            5' sequences, selection of clones, and sequence assembly.
FEATURES             Location/Qualifiers
     source          1..1232
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="19681"
                     94..1005
                     /codon_start=1
     CDS             94..1005
                     /codon_start=1
                     /product="SF2/ASF-like splicing modulator Srp30, putative"
                     /protein_id="AAM63132.1"
                     /db_xref="GI:21554051"
                     /translation="MSSRSSRTVYVGNLPGDIRERBVEDLFSKYGPVVQIDLKVPPRP
                     PGYAFVEFDDARDAEDAIHGRDGYDFDGHRLRVELAHGGRRSSDDTRGSFNGGGRGGG
                     RGRGDGGSRGGRGPSRRSEFRVLVTGLPSSASWQDLKDHMRKGGDVCFSQYRDARGTTGV
                     VDYTCYEDMKYALKKLDDTBFRNAFSNGYVRVREYDSRKDSRSPSRGSEYSKSRSRSR
                     GRSVSRSRSRSRSRSRSRSPAKSTSRSPGPRSKSRSPSPRRSRSRSRSPLP
                     SVQKEGSKSPSKPSPAKSPIHTRSPSR"
```

ORIGIN

Alignment Scores:

```
Pred. No.:                      Length:         1232
Score:            865.00        Matches:         184
Percent Similarity:   74.32%    Conservative:     33
Best Local Similarity: 63.01%   Mismatches:       45
Query Match:      59.61%        Indels:           30
DB:                             Gaps:              7
```

2.94e-38

US-10-014-927-19 (1-279) x AY085920 (1-1232)

```
Qy    1   MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys  20
              |||||||||||||||||||||||||||||||||||||||
```

RESULT 8

```
AY085920
LOCUS       AY085920        1232 bp    mRNA      linear    PLN 14-APR-2003
DEFINITION  Arabidopsis thaliana clone 19681 mRNA, complete sequence.
ACCESSION   AY085920
```

REFERENCE
AUTHORS   Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
           Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
           Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
           Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
           Ecker,J.R. and Theologis,A.
TITLE   Arabidopsis Full Length cDNA Clones
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 1523)
AUTHORS   Yamada,K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Lee,J.M.,
           Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y.,
           Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R.,
           Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
           Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C.,
           Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
           Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E.,
           Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE   Direct Submission
JOURNAL   Submitted (13-SEP-2001) Plant Gene Expression Center, 800 Buchanan
           Street, Albany, CA 94710, USA
REFERENCE 3 (bases 1 to 1523)
AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
           Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
           Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R.,
           Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J.,
           Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
           Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K.,
           Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE   Direct Submission
JOURNAL   Submitted (23-SEP-2002) Plant Gene Expression Center, 800 Buchanan
           Street, Albany, CA 94710, USA
COMMENT   RIKEN Genomic Sciences Center (GSC) members carried out the
           collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
           Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
           Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
           Hayashizaki,Y. and Shinozaki,K.
           The Salk, Stanford, PGEC (SSP) Consortium members carried out the
           sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M.,
           Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Quach,H.L.,
           Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G.,
           Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M.,
           Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W.,
           Ecker,J.R. and Theologis,A.
           Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
           this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
           contributed equally to this work as PIs.
FEATURES   Location/Qualifiers
     source   1..1523
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL08-15-J15 (R11393)"
                     /ecotype="Columbia"
                     /note="This clone is in a modified pBluescript vector
                     (FLC-1) as a BamHI/XhoI insert."
     gene   1..1523
                     /gene="At4g02430"
     misc_feature   1..1507
                     /gene="At4g02430"
                     /note="This may be a potential product of a partially
                     processed transcript corresponding to gene At4g02430."
                     /evidence=experimental
ORIGIN

Alignment Scores:
Pred. No.:   1.37e-37   Length: 1523
Score: 854.50   Matches: 181
Percent Similarity: 72.92%   Conservative: 29
Best Local Similarity: 62.85%   Mismatches: 42
Query Match: 58.89%   Indels: 36
DB: 8   Gaps: 6

Db   94 ATGAGACGAGTCGTTCGAGTAGAACCGTGTACGTCGGAAACCTTCCTGGCGATATCCGTGAG 153
           ||||||||||||||||||||||||||||||||||||||||||||||:|||
Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIle 40
AGAGAGGTCGAAGATTTGTTCAGTAAGTATGGACCCGTTGTTCAAATTGATTTGAAGGTT 213
           ||||||||:|||
Db 154 AGAGAGGTCGAAGATTTGTTCAGTAAGTATGGACCCGTTGTTCAAATTGATTTGAAGGTT 213
Qy   41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
           ||||:::|
Db 214 CCTCCAAGGCCCCCTGGTTATGCATTCGTTGAGTTTGATGAGCCTCGGGATGCTGAAGAT 273
Qy   61 AlaIleTyrGlyArgAspGlyTyrArgAspPheAspGlyCysArgLeuArgValGluIleAla 80
           |||||||||||||||||
Db 274 GCTATTCATGGTCGTGATGGCTATGACTTTGATGGGCATCGTTGAGGGTGGAATTGGCG 333
Qy   81 HisGlyGlyArgArgPheSerSerValAspArgTyrSerSerTyrSer------ 98
           |||||||||||||:::::|
Db 334 CATGGTGGGAGGCGT------TCATCAGTAGTACTCCGGGTAGTTTCAATGGTGGT 384
Qy   99 ---------AlaSerArgAlaProSerArgArgSer 107
                    ||||||||
Db 385 GGCCGTGGTGGTGGTCGCGGCCGTGGTGGATGGTGGTAGTCGTGGGCCATCTAGGAGATCA 444
Qy 108 AspTyrArgValLeuValThrGlyLeuProSerAlaSerTrpGlnAspLeuLysAsp 127
           :::||||||||||||
Db 445 GAGTTTGTGTTCTAGTCACAGGCTTGCCTTCATCTGCTTCTGGCAAGATCTCAGGAT 504
Qy 128 HisMetArgLysAlaGlyAspAlaValCysPheSerGluValPheProAspArgLysGlyMet 147
           |||||||||||||:::::|
Db 505 CACATCGTAAGAGGAGGCGATGTCTGTTTCTCGCAAGTGTACCGTGATGCTGAGGGGACA 564
Qy 148 SerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAsp 167
           :::||||||||:::||
Db 565 ACTGGAGTTGTTGATTACACCTGCTATGAGGACATGAAGTATGCGCTGAAAAAAGTCGAC 624
Qy 168 AlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSer 187
           |||||||||::::|
Db 625 GACACCAGAGTTTCGAAATGCGTTTCGAAATGGATATGTCCGGGTTAGAGATATGATTCA 684
Qy 188 ArgSerValSerArgSerArgProAspAspSerLysSer-----ArgSerArgSerSer 206
           |:::||
Db 685 AGGAAGGAGATTCTAGGAGTCCTAGCCGGGGAAGATCCTATTCTAAGAGCCGGCAGCAGC 744
Qy 207 ArgGlyProSerCysSerTyrSerSerArgSerValSerProAlaArgSerIle 226
           |||||:::||
Db 745 CGTGGACGAAGCGTG-------AGCCGAAGCAGGAAGCAGGAGCAGGAGCAGGACAGA 795
Qy 227 SerProArgSerArgProLeuSerArgArgSerArgSerLeuTyrSerSerValSerArgSer 246
           |:::::::|
Db 796 AGTCCCAAGGCAAAG-----TCTTCCACGTAGGTCCCCTGCAAAATCTACATCAAGATCT 849
Qy 247 GlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSer-----Arg 264
           ||||||||||
Db 850 CCTGGCCCCCGC-------------TCGAAGTCAAGTCACCGTCTCCAAGAAGA 891
Qy 265 SerArgSerArgSerAsnSerProValSerProVal 276
           ||||||:::||
Db 892 TCGCGTTCAAGATCAAGATCCTCTACCTTCTGTT 927

RESULT 9
AY056185
LOCUS   AY056185   1523 bp   mRNA   linear   PLN 23-SEP-2002
DEFINITION Arabidopsis thaliana At4g02430 mRNA sequence.
ACCESSION   AY056185
VERSION   AY056185.1 GI:15810292
KEYWORDS   FLI_cDNA.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 1523)
AUTHORS   Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
            Quach,H.L., Tang,C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C.,

```
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE  1
AUTHORS    Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
           Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
           Kawai,J., Hayashizaki,Y. and Shinozaki,K.
TITLE      Arabidopsis thaliana full-length cDNA
JOURNAL    Published Only in Database (2002)
REFERENCE  2  (bases 1 to 2059)
AUTHORS    Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
           Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
           Kawai,J., Hayashizaki,Y. and Shinozaki,K.
TITLE      Direct Submission
JOURNAL    Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
           Center; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
           230-0045, Japan (E-mail:mseki@gsc.riken.go.jp,
           URL:http://pfgweb.gsc.riken.go.jp, Tel:81-45-503-9625,
           Fax:81-45-503-9586)
COMMENT    An Arabidopsis full-length cDNA library was constructed essentially
           as reported previously (Seki et al. (1998) Plant J. 15:707-720;
           Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
           and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et
           al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
           This clone is in a modified pBluescript vector.
           Please visit our web site (http://pfgweb.gsc.riken.go.jp/) for
           further details.
FEATURES             Location/Qualifiers
     source          1..2059
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
                     /clone="RAFL19-32-J05"
                     /note="common name: thale cress"
     gene            1..2059
                     /gene="At4g02430/T14P8_21"
     CDS             153..689
                     /gene="At4g02430/T14P8_21"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="BAC42705.1"
                     /db_xref="GI:26451203"
                     /translation="MSSRSSRTIYRGNLPGDIRREVEDLFSKYGPVVQIDLKIPPRP
                     PGYAFVEFEDARDADDAIYGRDGYFDGHHLRVELAHGGRRSSHDARGSYSGRGRGGR
                     GGDGGGRRERGPSRSEYRVVVSGLPSSASWQDLKDHMRKGGEBVCFSQVPRDRGTTG
                     IVDYTSYEDMKYAIKKAR"
```

```
ORIGIN
Alignment Scores:
Pred. No.:                          2059
Score:                  854.50      Matches:        181
Percent Similarity:     72.92%      Conservative:    29
Best Local Similarity:  62.85%      Mismatches:      42
Query Match:            58.89%      Indels:          36
DB:                                 Gaps:             6

US-10-014-927-19 (1-279) x AK18074 (1-2059)

Qy    1 MetSerSerArgTrpAsnArgThrIleTyrArgGlyAsnLeuProGlyAspIleArgLys 20
          |||||||||||  :|||::|||||     :|||||||||||||||::|
Db  153 ATGAGCAGCCGTTCGAGTAGAACGATTTACGTCGGGAACCTTCCCGGCGATATCCGTGAA 212

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuAspLysIle 40
          |||||||||||||||:||||::|||::||||||||||||:||::||
Db  213 AGAGAAGTTGAAGACTTGTTCAGTAAGTATGGACCTGTTGTTCAAATCGATTTGAAGATT 272

Qy   41 ProProArgProProGlyTyrAlaPheValGluAspProArgAspAlaAspAsp 60
          ||||||
Db  273 CCGCCGAGCCCTCCAGGCTATGCCATTCGTGGAGATGCTCGTGATGCTGATGAT 332

Qy   61 AlaIleTyrGlyArgAspGlyTyrPheAspGlyCysArgLeuArgValGluIleAla 80
```

```
US-10-014-927-19 (1-279) x AY056185 (1-1523)

Qy    1 MetSerSerArgTrpAsnArgThrIleTyrArgGlyAsnLeuProGlyAspIleArgLys 20
Db   49 ATGAGCAGCCGTTCGAGTAGAACGATTTACGTCGGGAACCTTCCCGGCGATATCCGTGAA 108

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuAspLysIle 40
Db  109 AGAGAAGTTGAAGACTTGTTCAGTAAGTATGGACCTGTTGTTCAAATCGATTTGAAGATT 168

Qy   41 ProProArgProProGlyTyrAlaPheValGluAspProArgAspAlaAspAsp 60
Db  169 CCGCCGAGCCCTCCAGGCTATGCCATTCGTGGAGATGCTCGTGATGCTGATGAT 228

Qy   61 AlaIleTyrGlyArgAspGlyTyrPheAspGlyCysArgLeuArgValGluIleAla 80
Db  229 GCAATTTATGGCCGTGATGGTTATGACTTTGATGGGCATCATTACGGGTGGAACTAGCT 288

Qy   81 HisGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAla--- 99
Db  289 CATGTGTGGGAGGCGTTCATCACATGATGCA-------CGCGGTAGTTATATGTGGTCGT 339

Qy  100 ------SerArgAlaProSerArgArg 106
Db  340 GGTCGTGCGCGGTCGTGGTGGTGGTGGTGACCGGTGGTGGTCGTGAACGTGACCATCTAGGAGA 399

Qy  107 SerAspTyrArgGlyValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126
Db  400 TCAGACTACCGGTGTAGTGTCAGGTTTGCCTTCATCTGCGTCCTGGCAAGACCCTCAAG 459

Qy  127 AspHisMetArgLysAlaGlyAlaValCysPheSerGluValPheProProArgLysGly 146
Db  460 GATCACACATGCGTAAAGGCAGGAGCAGAAGTTTGTTTTCTCAAGTGTTCGTGATGGTAGGT 519

Qy  147 MetSerSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle-ArgLysIle 166
Db  520 ACAACTGGAATTGTAGATTATACCAGCTACGAGGACACTACGAAATATGCGGATAAAAAGCT 579

Qy  166 uAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGl 186
Db  580 CGATGACACAGAGTTTCGGAATGCGTTTTCTCATGAATATGTTCGGGTTAGAGAATATGA 639

Qy  186 uSerArgSerValSerArgSerProAspAspSerLysSerTyr--ArgSerArgSerAr 205
Db  640 TTCAAGGAGGGATTCGAGGAGCCCCAGCCGTGGAGATCCTATTCTAAGAGCCGCAGCCG 699

Qy  205 gSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSe 225
Db  700 TGGACGGAGTCCCAGC--------CGTAGTCGTAGCCCCAGCAGAAGCAGGAGCAAGAG 750

Qy  225 rIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerSerAr 245
Db  751 CAGAGTTCAAAGGCTAAATCTTTGCGTAGATGCCCTGCA------AAATCTACATCGAG 804

Qy  245 gSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArgSe 265
Db  805 ATCT------------------CCTCGCCTTGCGCTCCCCGCTC 828

Qy  265 rArgSerArgSerAsnSerPro 272
Db  829 TAAG*CGAGGT*CACTGTCTCCA 850
```

Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.S., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
Ecker,J.R.

TITLE　　Direct Submission
JOURNAL　Submitted (18-APR-2003) Salk Institute Genomic Analysis Laboratory
　　　　　(SIGnAL), Plant Biology Laboratory, The Salk Institute for
　　　　　Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
　　　　　USA

COMMENT　RIKEN Genomic Sciences Center (GSC) members carried out the
　　　　　collection and clustering of RAFL cDNAs (RAFL cDNA.: 'RIKEN
　　　　　Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
　　　　　Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
　　　　　Hayashizaki,Y. and Shinozaki,K.

　　　　　The Salk, Stanford, PGEC (SSP) Consortium members constructed and
　　　　　sequenced the pUNI (ORF) clones using the RAFL cDNAs: Shinn,P.,
　　　　　Chen,H., Cheuk,R., Kim,C.J., Bowser,L., Chan, M.M., Chang,C.M.,
　　　　　Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
　　　　　Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm,C.J.,
　　　　　Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
　　　　　Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and
　　　　　Ecker,J.R.

　　　　　Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
　　　　　this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
　　　　　contributed equally to this work as PIs.

FEATURES　　　　　Location/Qualifiers
　source　　　　　1..858
　　　　　　　　　　/organism="Arabidopsis thaliana"
　　　　　　　　　　/mol_type="mRNA"
　　　　　　　　　　/db_xref="taxon:3702"
　　　　　　　　　　/chromosome="1"
　　　　　　　　　　/clone="U15578"
　　　　　　　　　　/ecotype="Columbia"
　　　　　　　　　　/note="This clone is in pUNI 51"
　CDS　　　　　　1..858
　　　　　　　　　　/note=".SF2/ASF-like splicing modulator Srp30, putative'"
　　　　　　　　　　/codon_start=1
　　　　　　　　　　/product="Atig02840"
　　　　　　　　　　/protein_id="AAP13424.1"
　　　　　　　　　　/db_xref="GI:30023782"
　　　　　　　　　　/translation="MSSRSSRTVVYGNLPGDIRREREVEDLFSKYGPVVQIDLKVPPRP
　　　　　　　　　　PGYAFVEFDDARDAEDAIHGRDGYDFDGHLRVELAHGGRRSDDTRGSFNGGRGGG
　　　　　　　　　　RGRGDGSGRGPSRRSEFRVLVTGLPSSASWQDLKDHMRKGGDVCFSQYYRDARGTTGV
　　　　　　　　　　VDYTCYEDMKYALKKLDDTEFRNAFSNGYVRVREYDSRKDSRSPSRGRSYSKSRSRSR
　　　　　　　　　　GRSVSRSRSRSRSRSRSRSRSRSRPAKSTSRSPGPRSKSRSPSPRRWITVETLDHL
　　　　　　　　　　DHNIISGFL"

ORIGIN

Alignment Scores:
Pred. No.:　　　　　　1.82e-37　　　　　Length:　　　　　858
Score:　　　　　　　　847.00　　　　　Matches:　　　　177
Percent Similarity:　　73.08%　　　　Conservative:　　32
Best Local Similarity:　61.89%　　　　Mismatches:　　　41
Query Match:　　　　　58.37%　　　　Indels:　　　　　36
DB:　　　　　　　　　　　　　　　　　Gaps:　　　　　　6

US-10-014-927-19 (1-279) x BT006316 (1-858)

Qy　　　1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
　　　　　|| || | | || ||| |||| | | ||||||| || |:|
Db　　　1 ATGAGAGTCGTTCAGAGACCGGTAGTAGACCGGTACGTCGGAAACCTTCCTGGCGATATCCCGTAGCGAG 60

Qy　　21 CysGluValGluAspPheTyrLysTyrGlyProIleValAspIleArgLeuLysIle 40
　　　　　||| | || | | ||| | || | ||||||||| |||||:|:
Db　　61 AGAGAGGTCGAAGATTTGTTCAGTAAGTATGGACCTGTTGTTCAATTGATTTGAAGGTT 120

Qy　　41 ProProArgProProGlyTyrAlaPheValGluPheAspAspAlaArgAspAlaAspAsp 60
　　　　　|| || || | | |: || || :||:||:|| || || |:
Db　121 CCTCCAAGGCCTCCTGGTTATGCAATTCGTTGATGAGTTTGATGATGCCTGGAAGAT 180

LOCUS　　　　　BT006316　　　　　858 bp　　　mRNA　　　linear　　PLN 18-APR-2003
DEFINITION　　Arabidopsis thaliana Atig02840 mRNA, complete cds.
ACCESSION　　BT006316
VERSION　　　BT006316.1　　GI:30022781
KEYWORDS　　FLI_cDNA.
SOURCE　　　Arabidopsis thaliana (thale cress)
　ORGANISM　Arabidopsis thaliana
　　　　　　Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
　　　　　　Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
　　　　　　rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE　　1 (bases 1 to 858)
　AUTHORS　　Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Bowser,L., Carninci,P.,
　　　　　　Chan,M.M., Cheng,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
　　　　　　Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
　　　　　　Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
　　　　　　Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
　　　　　　Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
　　　　　　Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
　　　　　　Ecker,J.R.
　TITLE　　Arabidopsis ORF clones
　JOURNAL　Unpublished
REFERENCE　　2 (bases 1 to 858)
　AUTHORS　　Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Bowser,L., Carninci,P.,
　　　　　　Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,

Db　　333 GCAATTTATGGCCGTGATGGTTATGACTTTGATGGGCATCATTACGGGTGGAACTAGCT 392

Qy　　81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerAla--- 99
　　　　　||| | ||| ||
Db　　393 CATGGTGGAGGCGTTCATCACATGATGCA-------CGCGGTAGTTATAGTGGTCGT 443

Qy　100 -------------------SerArgAlaProSerArgArg 106
Db　　444 GGTCGTGCGGTCGTGGTGGTGGTGACGGTGGTCGTGAACGTGGACCATCTAGGAGA 503

Qy　107 SerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126
Db　　504 TCAGAGTACCGGCGTTGAGTGTCAGGTTGCCTTCATCTGCGTCCTGGCAAGACCTCAAG 563

Qy　127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146
Db　　564 GATCACATGCGTAAAGGAGGAGAAGTTTGTTTTTCCAAGTGTTTCGTGATGGTAGAGGT 623

Qy　147 MetSerGlyValValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle-ArgLysLe 166
Db　　624 ACAACTGGAATTGTAGATTATACCAGCTACGAGGACATGAAATATCGGATAAAAAAAGCT 683

Qy　166 uAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrG1 186
Db　　684 CGATGACACAGAGTTTCGGAATGCGTTTTCTCATGAATATGTTCGGGTTAGAGAATATGA 743

Qy　186 uSerArgSerValSerArgSerProAspSerLysSerTyr--ArgSerArgSerAr 205
Db　　744 TTCAAGAGGAGGATTCGAGGAGCCCCAGCCGTGGAAGATCCTATTCTAAGAGCCGCAGCCG 803

Qy　205 gSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSe 225
Db　　804 TGGACGGAGTCCCAGC--------CGTAGTCGTAGCCGCAGCAGAAGCAGGAGCAAGAG 854

Qy　225 rIleSerProArgSerArgProLeuSerArgSerSerArgSerLeuTyrSerSerValSerAr 245
Db　　855 CAGAAGTCCAAAGGCTAAATCTTTGCGTAGATCGCCTGCA------AAATCTACATCGAG 908

Qy　245 gSerGlySerLeuArgArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArgSe 265
Db　　909 ATCT----------------------------CCTCGCTTCTCGCTCCCCGCTC 932

Qy　265 rArgSerArgSerAsnSerPro 272
Db　　933 TAAGTCGAGGTCACTGTCTCCA 954

COMMENT

e-mail for correspondence: arab@sequence.stanford.edu

The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Tripp,M.,
Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R.,
Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H.,
Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Tripp,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES   Location/Qualifiers
    source   1..1540
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="RAFL04-19-E10 (R15578)"
                     /ecotype="Columbia"
                     /note="This clone is in pBluescript"
    gene   1..1540
                     /gene="At1g02840"
                     /note="synonym: F22D16.16"
    CDS   49..906
                     /gene="At1g02840"
                     /codon_start=1
                     /product="SF2/ASF-like splicing modulator Srp30, putative"
                     /protein_id="AAM91541.1"
                     /db_xref="GI:22135918"
                     /translation="MSSRSSRTVYVGNLPGDIREREVEDLFSKYGPVVQIDLKVPPRP
                     PGYAFVEFDDARDAEDAIHGRDGYDFDGHRLRVELAHGGRRSSDDTRGSFNGGRGGG
                     RGRDGGSGRGPSRRSEFRVLVTGLPSSASWQDLKDHMRKGGDVCFSQVYRDARGTTGV
                     VDVTCYEDMKYALKKLDDTEFRNAFSNGYVRVREYDSRKDSRSRGRSYSKSRSRSR
                     GRSVSRSRSRSRSRSRSRSPKAKSSRSRSPGPRSKSRSPSPRRWITVETLDHL
                     DHNIISGFL"

ORIGIN

Alignment Scores:
Pred. No.:   Length: 1540
Score: 3.52e-37   Matches: 177
Percent Similarity: 847.00   Conservative: 32
Best Local Similarity: 73.08%   Mismatches: 41
Query Match: 61.89%   Indels: 36
DB: 58.37%   Gaps: 6
                            8

US-10-014-927-19 (1-279) x AY128338 (1-1540)

Qy   1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   20
            |||||||||||| ::||||||||||||||||||||||||||||||||||||||||
Db   49 ATGAGCAGTCGTTCAGTAGACTAGAACCGGTCACGTCGGAAACCTTCCTGGCGATATCCGTGAG 108

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuAspLeuLysIle   40
            |||||||||||||||:||| :||||||| :|:|| |||||||||||:||||||||||||
Db 109 AGAGAGGTCGAAGATTTGTTCAGTAAGTAAGTGACCTGTTGTCAATTGATTTGAAGGTT 168

Qy   41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp   60
            |||||||||||| ||:|| |||||||||||| ||||||||||||||||||||:||
Db 169 CCTCCAAGGCCTCCTGGTTATGCCATTCGTTCGTGAGTTGATGATGCTGAAGAT 228

Qy   61 AlaIleTyrArgAspAspPheAspGlyTyrAspPheCysArgLeuArgValGluIleAla   80
            |||||||||||| ||| :|||||| ::|||| |||||||||||||:||||||||||
Db 229 GCTATTCATGTCGTGATGGCTATGACTTTGATGGGCATCGTTTGAGGGGTTGGAATTGGCG 288

Qy   81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSer------   98
            :||||:| :||||||||||| |:|||| :| :||||||||||
Db 289 CATGGTTGGGAGGCGT----------TCATCAGATGACTACCTCGGGGTAGTTTCAATGGTGGT 339

```
Qy        99 ---------------------------------------------AlaSerArgAlaProSerArgArgSer 107
                                                        |||||||||||||||||
Db       340 GGCCGTGGTGGTGGTCGCGGCCGTGTGATGGTGGTAGTCGTGGGCCATCTAGGAGATCA 399

Qy       108 AspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAsp 127
             :::|||||:||||| ||||||||||:|||||||||||||:::
Db       400 GAGTTTCGTGTTCTAGTCACAGGCTTGCCTTCATCGGCTTCTTGGCAAGATCTCAAGGAT 459

Qy       128 HisMetArgLysAlaGlyAspValCysPheSerGluValPheProArgArgLysGlyMet 147
             ||||| |||| |||||||||| ::|:||  :||:|||  :::|||
Db       460 CACATGCGTAAAGGAGGCGGATGTCGTTTCCGCAAGTGTACCGTGATGCTAGAGGGACA 519

Qy       148 SerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgGlyLeuAsp 167
             :::|:|||||||||::::| |||:||::||||| |||||:||:||||::||:||
Db       520 ACTGGAGTTGTTGATTACACCTGCTATGAGGACATGAAGTATGCCTGAAAAAGCTCGAC 579

Qy       168 AlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSer 187
             ||||||||||:|||||||||::||| |||||
Db       580 GACACAGAGTTTCGAAATGCGTTTCGAATCGATATGTCCCGGTTAGAGATATCGATTCA 639

Qy       188 ArgSerValSerArgSerProAspAspSerSerTyr---ArgSerArgArgSerSer 206
             ||||||||| :::||||  :::|||
Db       640 AGGAAGGATTCTAGAGTCCTAGCCGGGAAGATCCTATTCTAAGAGCCGCAGCCGCAGC 699

Qy       207 ArgGlyProSerCysSerTyrSerSerArgSerValSerProAlaArgSerIle 226
             ||:::||||| :::|||  :::|||
Db       700 CGTGGACGAAGCGTG------AGCCGAAGCAGGAGCAAGCAGGAGCAGGACGAGCAGA 750

Qy       227 SerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSer 246
             |||||||||||:::|||
Db       751 AGTCCCAAGGCAAAG-----TCTTCACGTAGTCCCCTGCAAAATCTACATCAAGATCT 804

Qy       247 GlySerLeuLeuArgGlyAlaGlyAspTrpIleSerGlnSerArgSerLysSerSerArgArg 266
                                            ||:::|||
Db       805 CCTGGC-----------------------------------CCCCGTCTCGAAG 822

Qy       267 SerArgSerAsnSerPro 272
             ||||||:::|||
Db       823 TCAAGGTCACCGTCTCCA 840
```

```
RESULT 13
AK071503
AK071503
LOCUS         AK071503               1293 bp    mRNA     linear   PLN 24-JUL-2003
DEFINITION    Oryza sativa (japonica cultivar-group) cDNA clone:J023091N08, full
              insert sequence.
ACCESSION     AK071503
VERSION       AK071503.1  GI:32981526
KEYWORDS      FLI_CDNA; CAP trapper.
SOURCE        Oryza sativa (japonica cultivar-group)
ORGANISM      Oryza sativa (japonica cultivar-group)
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzeae; Oryza.
REFERENCE     1
AUTHORS       The Rice Full-Length cDNA Consortium, National Institute of
              Agrobiological Sciences Rice Full-Length cDNA Project Team;
              Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
              Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
              Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
              Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,
              Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
              Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
              Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
              Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
              Yoshino,M. and Hayashizaki,Y.
TITLE         Collection, mapping, and annotation of over 28,000 cDNA clones from
```

```
JOURNAL    japonica rice
           Science 301 (5631), 376-379 (2003)
MEDLINE    22752273
PUBMED     12869764
REFERENCE  2  (bases 1 to 1293)
AUTHORS    Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
           Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
           Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
           Hori,F., Hotta,I., Iida,J., Iida,Y., Ishii,Y., Imamura,K.,
           Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
           Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
           Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.
           Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
           Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
           Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
           Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
           Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
           Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
           Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
           Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
           Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
           Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,T., Tagami-Takeda,Y.,
           Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
           Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
           Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
           Yoshimura,A.
TITLE      Direct Submission
JOURNAL    Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
           Agrobiological Sciences, Department of Molecular Genetics, Head of
           Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
           305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
           Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT    This clone is one of the 28K full-length cDNA clones from japonica
           rice.
           URL : http://cdna01.dna.affrc.go.jp/cDNA/
           NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
           Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
           Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
           Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
           Yamamoto,M.
           FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
           Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
           Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,T., Miura,J.,
           Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
           Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
           Yoshimura,A., Matsubara,K. and Murakami,K.
           Genome Exploration Research Group in Riken Genomic Sciences Center
           and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
           Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
           Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
           Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
           Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
           Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
           Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Osato,N.,
           Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
           Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
           Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
           Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
           Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
           Yasunishi,A. and Hayashizaki,Y.
FEATURES            Location/Qualifiers
     source         1..1293
                    /organism="Oryza sativa (japonica cultivar-group)"
                    /mol_type="mRNA"
                    /cultivar="Nipponbare"
                    /db_xref="taxon:39947"
                    /clone="J023091N08"
ORIGIN

Alignment Scores:
Pred. No.:      1293
Score:          3.27e-37         Length:       1293
Percent Similarity: 75.27%       Matches:      846.00
                                 Conservative: 24
```

```
Best Local Similarity:  66.67%      Mismatches:  44
Query Match:            58.30%      Indels:      26
DB:                         8       Gaps:         7

US-10-014-927-19 (1-279) x AK071503 (1-1293)

Qy      1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
          ||||| ||| |||||  ||||||  ::::
Db     78 ATGAGCCACGCAACAGCCGCCACCATCATGTGGGCCAATCTCCCCGAGACATTCGCGAG 137

Qy     21 CysGluValGluAspPheTyrLysTyrGlyTyrGlyProIleValAspLeuAspLysIle 40
          |||||||||||||||||||||||||| |||::
Db    138 AGGGAAGTGGAGGATCTCTTCTATAAGTATGGCGTATTGTGGATATTGACTTGAAAATT 197

Qy     41 ProProArgProProGlyTyrAlaPheValAlaPheGluGluAspProArgAspAlaAspAsp 60
          |||||||||||||||||
Db    198 CCTCCAAGGCCTCCTGGTTATGCTTTTGTTGAGTTTGAAGACCCTGTGATGCACAAGAT 257

Qy     61 AlaIleTyrGlyArgGlyTyrArgAspPheAspGlyCysArgGluArgValGlulleAla 80
          |||||| |||||||||::||||
Db    258 GCAATTTATGGCCGTCATGGATAGACTTTGATGGCCACAGTTGCGGTGAGTTAGCT 317

Qy     81 HisGlyArgArgPheSerProSerValAspArgTyrSerSerAlaSer 100
          ||||||||||||||||:::::::
Db    318 CATGTGGCAGA------GGCCCTTCTTTGATCGA---TCAAGTAGCTATGCAGTGCT 368

Qy    101 ------ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProPro 118
          ::::::||||||||||||||||||
Db    369 GGGGCGCCGGTGGTGCCGCTAAACGTACTGATTACCGAGTTAGTTACTGGATTACCTTCG 428

Qy    119 SerAlaSerTrpGlnAspLeuLysMetArgLysAlaGlyAspValCysPheSer 138
          ||||||||||||||:::::||||||
Db    429 TCAGCACATCATGGCAAGACCTCAAGGATCATATGAGGGAGCTGGTGATGTCTGTTTCTCT 488

Qy    139 GluValPheProArgArgGlyMetSerGlyValValAspTyrSerAsnTyrAspAsp 158
          :::||||||| ::: ||||||||||::::||||||
Db    489 GATGTTTTACCGTGAG--GGTGGAGC-AATGGAATTGTGATTATACCAACATGAGGAT 544

Qy    159 MetLysTyrAlaIleLysLeuAspAspThrGluPheArgAsnAlaPheSerSerAla 178
          ||||||  ||||||||||||  |||  ||||||||
Db    545 ATGAAACAAGCGATAAGAAAGCTTGATGACTCAGAGTTCCGTAATGCATTTTCACGGGCA 604

Qy    179 TyrIleArgValArgGluTyrGluSerArgSerArgSerProAspAspSerSer 198
          ||||||  ||||||||||||||:::|||
Db    605 TATATCAGGGTGAGGGAGTATATAGTCTGCTAAGCCAGCCGCTCTCGTTCCAGAGGTAGA 664

Qy    199 SerTyr------------ArgTyrArgSerArgSerArgGlyProSerCysSerTyr 213
          :::||::               |||||||||||||
Db    665 AGCCACAGCCGCGTCTATCCTCAGAAGCAGAAGCCGCCGCCTCG--------------- 709

Qy    214 SerSerLysSerArgSerValSerProAlaArgArgSerIleSerProArgSerProLeu 233
          :::||||||||  ::|||||||||||
Db    710 ATCTCCAAGAGCAGAAGCCGCAGTTATGCAGAAGCAGGAGTCCAAGATCTAGGTCTGCT 769

Qy    234 SerArgSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuArgArgAlaGly 253
          |||||||||
Db    770 TCTCGGTCTCACTCACCTGTGAAAGAAAGATCTAGGTCGGCT------------------ 811

Qy    254 AspTrpIleSerArgGlnSerArgLysSerArgSerArgSerArgSerAsnSerPro 272
          ::: :: |||||
Db    812 -------TCACAGTCTCGCTCACCTGGCAAAGGAAAAGATCACCAATCAAGATCCCCA 859
```

RESULT 14
ATHRIBNUCP
LOCUS   ATHRIBNUCP   1160 bp   mRNA   linear   PLN 27-APR-1993
DEFINITION   Arabidopsis thaliana ribonucleoprotein mRNA, complete cds.
ACCESSION   M98340
VERSION   M98340.1 GI:166843
KEYWORDS   homologue; ribonucleoprotein; splicing factor; splicing factor
               SF-2.
SOURCE   Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1160)

REFERENCE
AUTHORS   Lazar,G.
JOURNAL   Unpublished (1992)
COMMENT   Original source text: Arabidopsis thaliana (strain Landsberg
             erecta) (library: Lambda gt11) root cDNA to mRNA.

FEATURES   Location/Qualifiers
source   1..1160
             /organism="Arabidopsis thaliana"
             /mol_type="mRNA"
             /strain="Landsberg erecta"
             /db_xref="taxon:3702"
             /tissue_type="root"
             /tissue_lib="Lambda gt11"
CDS   31..942
             /note="human splicing factor SF-2 homologue"
             /codon_start=1
             /product="ribonucleoprotein"
             /protein_id="AAA32856.1"
             /db_xref="GI:166844"
             /translation="MSSRSRTVVGNLPGDIREREVEDLFSKYGPVVQIDLKVPPRP
             PGYAFVEFDDARDADAIHGRDGYDFDGHRLRVELAHGGRRSSDDTRGSPNGGRGGG
             RGRGDGGRSGPSRRSEFRVLVTWLASSASWQDLKDHIAKGGDVCFSQVYRDARGTTGV
             VDYTCYEDMKYALKKLDDTEFRNAFSNGYVRVREYDSRSRGRSYSKSSRSR
             GRSVSRSRSRSRSRSRPKAKSRRSPAKSTRSPGPRSKSRSPPRSRSRSRSPLP
             SVQKEGSKSPSKPSPAKSPIHTRSPSR"

ORIGIN

Alignment Scores:
Pred. No.:   Length: 1160
Score: 6.88e-37   Matches: 180
Percent Similarity: 839.00   Conservative: 34
Best Local Similarity: 73.29%   Mismatches: 48
Query Match: 61.64%   Indels: 30
DB: 57.82%   Gaps: 7
                         8

US-10-014-927-19 (1-279) x ATHRIBNUCP (1-1160)

```
Qy      1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
          ||||| ||| |||||  ||||||  ::::
Db     31 ATGAGCAGTCGTTCGAGTAGAACCGTACGTCGGGAAACCTTCCTGGCGATATCCGTGAG 90

Qy     21 CysGluValGluAspPheTyrLysTyrGlyProIleValAspLeuAspLysIle 40
          ||||||||||||||||||||||::|||:|:
Db     91 AGAGAGGTCGAAGATTTGTTCAGTAGTAATGGACCGTGTGTTCAATTGATTGAGGTT 150

Qy     41 ProProArgProProGlyTyrAlaPheValGluPheAspAspAlaArgAspAlaAspAsp 60
          |||||||||||||||| |||||
Db    151 CCTCCAAGGCCCTCGTTACGCTTTTGTTGAGTTTGATGATGCATGGGGGATGCTGAAGAT 210

Qy     61 AlaIleTyrArgArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGlulleAla 80
          ||||||||||||||::||||||:
Db    211 GCTATTCATGGTCGTGATGGCTATGACTTTGATGGGCATCGTTTGAGGGGTGGAATTGCCG 270

Qy     81 HisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMet 98
          ||| :::|||  |||||||||||::: |||
Db    271 CACATGCCTAAAGGAGCTAGGAGGCGATGTCGTTTCTCCGCAAGTACCGTAGAGGGACA 501

Qy    148 SerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAsp 167
          :::||||||||||||||||||||:::::
```

Db 502 ACTGGAGTTGTTGATTACACCTGCTATGAGGACATGAAGTATGGCTGAAAAAGCTCGAC 561

Qy 168 AlaThrGluPheAlaPheSerLeuAlaTyrIleArgValArgGluTyrGluSer 187
        |||||||||||||| ||||||| :::::

Db 562 GACACAGAGTTTCGAAATGCGTTTTCGAATCGATATGTCCCGGTTAGAGAATATGATTCA 621

Qy 188 ArgSerValSerArgSerProAspAspSerLysSerTyr---ArgSerArgSerArgSer 206
              :::||| :::|||||||||

Db 622 AGGAAGGATTCCAGGAGTCCTACGCCGGGAGATCCTATTCTAAGAGCCCGCCAGCCGCAGC 681

Qy 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIle 226
           ||||| ||:|| :::|||||||

Db 682 CGTGGACGAAGCGTG-------AGCCGAAGCAGGAGCAGGAGCAGGAGCAGA 732

Qy 227 SerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSer 246
         ||||| :::::::::

Db 733 AGTCCCAAGGCAAAG-----TCTTCACGTAGGTCCCCTGCAAAATCTACATCAAGATCT 786

Qy 247 GlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSer------Arg 264
                |||:|||||||

Db 787 CCTGGCCCCGC-----------------TCGAAGTCAAGGTCACCGTGTCCCAAGAAGA 828

Qy 265 SerArgSerArgSerAsnSerProValSerProVal 276
         ||||||||| |||||||

Db 829 TCGCGTTCAAGATCAAGATCTCCTCCTTACCTTCTGTT 864

RESULT 15

AK106176
AK106176

LOCUS   AK106176   1352 bp   mRNA   linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:001-208-C08, full
           insert sequence.
ACCESSION   AK106176
VERSION   AK106176.1 GI:32991385
KEYWORDS   FLI_CDNA; oligo capping.
SOURCE   Oryza sativa (japonica cultivar-group)
  ORGANISM Oryza sativa (japonica cultivar-group)
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
           Agrobiological Sciences Rice Full-Length cDNA Project Team:,
           Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
           Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
           Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
           Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
           Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
           Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
           Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
           Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
           Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
           Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
           Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
           Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
           Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
           Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
           Yoshino,M. and Hayashizaki,Y.
  TITLE   Collection, mapping, and annotation of over 28,000 cDNA clones from
           japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
  MEDLINE   22752273
   PUBMED   12869764
REFERENCE   2 (bases 1 to 1352)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
           Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
           Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
           Hori,F., Hotta,I., Iida,J., Iida,J., Ishii,Y., Ikeda,R., Imamura,K.,
           Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
           Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
           Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
           Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
           Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
           Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
           Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
           Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,M.,
           Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
           Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
           Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
           Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
           Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
           Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
           Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
           Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
           Yoshimura,A.
  TITLE   Direct Submission
  JOURNAL   Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
           Agrobiological Sciences, Department of Molecular Genetics, Head of
           Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
           305-8602, Japan (B-mail:skikuchi@nias.affrc.go.jp,
           Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT   This clone is one of the 28K full-length cDNA clones from japonica
           rice.
           URL : http://cdna01.dna.affrc.go.jp/cDNA/
           NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K.,
           Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
           Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
           Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
           Yamamoto,M.
           FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
           Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
           Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
           Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
           Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
           Yoshimura,A., Matsubara,K. and Murakami,K.
           Genome Exploration Research Group in Riken Genomic Sciences Center
           and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
           Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
           Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
           Hiraoka,T., Hori,F., Ida,J., Imamura,K., Imotani,K., Ishii,Y.,
           Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
           Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
           Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
           Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
           Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
           Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
           Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
           Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
           Yasunishi,A. and Hayashizaki,Y.
FEATURES   Location/Qualifiers
    source   1..1352
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="mRNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="001-208-C08"
ORIGIN

Alignment Scores:
Pred. No.:   1.83e-36   Length: 1352
Score: 832.50   Matches: 184
Percent Similarity: 72.11%   Conservative: 28
Best Local Similarity: 62.59%   Mismatches: 57
Query Match: 57.37%   Indels: 25
DB: Gaps: 7

US-10-014-927-19 (1-279) x AK106176 (1-1352)

Qy   1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
        |||||| :||:|| :::|:|||||||||||||| :::
Db 145 ATGAGTAGGGCTGGACGAGCACAATTTATGTCGGGAACCTCCCGGGCCGACATCAGGGAG 204

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuAspLysLysIle 40
         |||||

Db 205 AGGGAGGTAGAGGATCGTGTTCTCACAAGTATCGACGCCATTGTTGATATTGACTTGAAGATC 264

Qy   41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgArgAlaAspAsp 60

```
Db  265 CCCCCAAGGCCCGCCTGGTTTGCTTTTGTTCTTGAGTTCGAAGATCCCCGTGATGCTGAAGAT  324

Qy   61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla      80
Db  325 GCAATCCGTGGACGGGGATGGCTATAATTTGATGGGAACCGTTTGAGGGTGGAACTTGCA      384

Qy   81 HisGlyGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer     100
Db  385 CATGGTGGGAGGAGGTAACTCTTCCTCC--------TTTAACAATTCTGGTGGAGGGGGA      435

Qy  101 Arg------AlaProSerArgSerArgTyrArgValLeuValThrArgGlyLeuProPro     118
Db  436 CGTCGTGGGGGGTGTATCTAGGCATACAGAGTATCGCCGTTCTGGTTACTCGGGGACTACCTTCT  495

Qy  119 SerAlaSerTrpGlnAspLeuLysAspHisMetArgGlyAlaGlyAspValCysPheSer     138
Db  496 TCAGCATCATGGCAAGACCTAAAGGACCATATGCGAAATGCTGGTGATGTTTGTTACTCT     555

Qy  139 GluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrArgAspAsp  158
Db  556 GAAGTGTACCGTGAGGGTGGTGTACTATAGGAATTGTAGACTACACAAACTATGATGAC     615

Qy  159 MetLysTyrAlaIleArgLysLeuAspAlaThrGlnPheArgAsnAlaPheSerSerAla    178
Db  616 ATGAAGTATGCATCAGGAAACTTGATGATTCTGAATTAAGAAATGCCTTCTCAAGGCC      675

Qy  179 TyrIleArgValAlaArgGluTyrGlu--------SerArgSerValSerArgSerProAsp  195
Db  676 TATATAAGGGTGAAGGAGTAGATGATGGAAAGCCAGCCGTTCATACTCACGAAGCCGAAGC    735

Qy  196 AspSer-----LysSerTyr---ArgSerArgSerArgGlyProSerCysSer          212
Db  736 CGAAGTCGAGGTGCCAGCTACAGCAGAAGCAGAGTCCAAGCAAGTCTCCCAAGGGAAAG     795

Qy  213 TyrSerSerLysSerArgSerValSerProAlaArgSerIleSerProArgSerArgPro    232
Db  796 TCATCACGGCCGTTCAGCATCTAGATCTCGATCGAGATCAGCTTCTTCCCGTTCAAGGATCA    855

Qy  233 LeuSerArgSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuLeuArgAla    252
Db  856 GAGTCAAAGGGGGGCCTCTTCCATCAAGATCACCAGCAAGATCTCAGTCTCCAAATACTTCT   915

Qy  253 --------GlyAspTrpIleSer--------------------------Gln          258
Db  916 CCTGCCAATGGTCAGTGATGCAGCAAGCCCCAGAAACGCAGCCCAAGCAGGAGCCCCCCAAG   975

Qy  259 SerArgSerLysSerArgSerArgSerArgSerAsnSerPro                     272
Db  976 AAACGTAGCCCCAAGCAGGAGCCCATCACGTTCCCGATCTCCT  1017
```

Search completed: March 26, 2004, 19:37:55
Job time : 4277 secs

OM protein - nucleic search, using frame_plus_p2n model

```
              Copyright  GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:       March 26, 2004, 17:13:02 ; Search time 491 Seconds
                              (without alignments)
                              2413.946 Million cell updates/sec

Title:        US-10-014-927-19
Perfect score: 1451
Sequence:     1 MSSRWNRTIYVGNLPGDIRK.........RSKSRSRSRSNSPVSPVISG 279

Scoring table: BLOSUM62
              Xgapop 10.0 , Xgapext   0.5
              Ygapop 10.0 , Ygapext   0.5
              Fgapop  6.0 , Fgapext   7.0
              Delop   6.0 , Delext    7.0

Searched:     3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:    6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10014927/runat_24032004_152925_9886/app_query.fasta_1_455
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10014927 @CGN_1_1_470 @runat_24032004_152925_9886 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :    N_Geneseq_29Jan04:*
          1:  geneseqn1980s:*
          2:  geneseqn1990s:*
          3:  geneseqn2000s:*
          4:  geneseqn2001as:*
          5:  geneseqn2001bs:*
          6:  geneseqn2002s:*
          7:  geneseqn2003as:*
          8:  geneseqn2003bs:*
          9:  geneseqn2003cs:*
          10: geneseqn2004s:*

              SUMMARIES
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.		Score	% Match	Query Length	DB	ID	Description
1		1359.5	93.7	925	3	AAC49248	Aac49248 Arabidops
2		1359.5	93.7	1213	3	AAC40180	Aac40180 Arabidops
3		1343	92.6	771	3	AAC49242	Aac49242 Arabidops
4		1272	87.7	762	6	ABZ13394	Abz13394 Arabidops
5		1159	79.9	714	3	AAC42865	Aac42865 Arabidops
6		875	60.3	4044	3	AAC81899	Aac81899 A. thalia
7		865	59.6	1232	3	AAC34814	Aac34814 Arabidops
8		863	59.5	1190	3	AAC38528	Aac38528 Arabidops
9		838.5	57.8	885	3	AAC42654	Aac42654 Arabidops
10		805	55.5	1241	3	AAC37026	Aac37026 Arabidops
11		584.5	40.3	2765	9	ADD22452	Add22452 HLA-B46 T
12		582.5	40.1	1428	4	AAH02904	Aah02904 Human she
13		582.5	40.1	1428	7	AAL60055	Aal60055 Human Pc0
14		550	37.9	1478	6	ABL29431	Abl29431 Drosophil
15		543	37.4	1584	6	ABI99368	Abi99368 Mouse isc
16	c	541.5	37.3	5285	2	AAV69915	Aav69915 Arabidops
17		537.5	37.0	498	5	AAC44527	Aac44527 Zea mays
18		517	35.6	1069	6	ABN96845	Abn96845 Gene #334
19		517	35.6	1162	5	AAF16226	Aaf16226 Human pro
20		517	35.6	1218	5	ABV30279	Abv30279 Human pro
21	c	517	35.6	1475	6	ABQ54393	Abq54393 Human ova
22		472	32.5	479	5	AAC35763	Aac35763 Zea mays
23		471	32.5	495	6	ABI99907	Abi99907 Mouse isc
24		457.5	31.5	568	3	AAC34694	Aac34694 Arabidops
25		456	31.4	3948	4	ABL29430	Abl29430 Drosophil
26		440.5	30.4	3697	7	ABX62942	Abx62942 Human act
27		436	30.0	289	7	AAC49252	Aac49252 Arabidops
28	c	432.5	29.8	545	7	ABX57250	Abx57250 Arabidops
29		431	29.7	1644	5	ABL30487	Abl30487 Drosophil
30		431	29.7	2076	6	ABK83839	Abk83839 Human cDN
31		418.5	28.8	1454	6	ABK84288	Abk84288 Human cDN
32		418.5	28.8	1534	7	ABX63268	Abx63268 Human cDN
33		416.5	28.7	2212	3	AAC98914	Aac98914 Human pan
34		412.5	28.4	759	6	ABS77425	Abs77425 Frog embr
35		407.5	27.7	446	6	ABV94679	Abv94679 Human pan
36		402.5	27.7	727	5	ABV16009	Abv16009 Human pro
37		391	26.9	492	3	AAC47771	Aac47771 Zea mays
38		388	26.7	345	6	ABQ85913	Abq85913 Arabidops
39		376	25.9	3872	5	AAS64831	Aas64831 DNA encod
40		374.5	25.8	471	3	AAC41641	Aac41641 Arabidops
41		374	25.8	444	7	ABV45807	Abv45807 Human pro
42		344	23.7	746	6	ABS77297	Abs77297 Frog embr
43		325	22.4	489	8	ACH35932	Ach35932 Human end
44		298	20.5	1110	9	ADD01227	Add01227 Human nuc
45		297	20.5	2367	2	AAZ41956	Aaz41956 Human myo

ALIGNMENTS

```
RESULT 1
AAC49248
ID   AAC49248 standard; DNA; 925 BP.
XX
AC   AAC49248;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 60471.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;  99US-0121825P.
PR   05-MAR-1999;  99US-0123180P.
PR   09-MAR-1999;  99US-0123548P.
PR   23-MAR-1999;  99US-0125788P.
PR   25-MAR-1999;  99US-0126264P.
PR   29-MAR-1999;  99US-0126785P.
PR   01-APR-1999;  99US-0127462P.
PR   06-APR-1999;  99US-0128234P.
PR   08-APR-1999;  99US-0128714P.
PR   16-APR-1999;  99US-0129845P.
```

us-10-014-927-19.rng

Mon Mar 29 09:55:42 2004

PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 09-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142205P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145219P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147003P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 09-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 10-AUG-1999; 99US-0147493P.
PR 11-AUG-1999; 99US-0147935P.
PR 12-AUG-1999; 99US-0148171P.
PR 13-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 16-AUG-1999; 99US-0148565P.
PR 17-AUG-1999; 99US-0148684P.
PR 18-AUG-1999; 99US-0149368P.
PR 20-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 23-AUG-1999; 99US-0149722P.
PR 25-AUG-1999; 99US-0149723P.
PR 26-AUG-1999; 99US-0149929P.
PR 27-AUG-1999; 99US-0149902P.
PR 27-AUG-1999; 99US-0149930P.
PR 30-AUG-1999; 99US-0150566P.
PR 31-AUG-1999; 99US-0150884P.
PR 01-SEP-1999; 99US-0151065P.
PR 07-SEP-1999; 99US-0151066P.
PR 10-SEP-1999; 99US-0151080P.
PR 13-SEP-1999; 99US-0151303P.
PR 15-SEP-1999; 99US-0151438P.
PR 16-SEP-1999; 99US-0151930P.
PR 20-SEP-1999; 99US-0152363P.
PR 22-SEP-1999; 99US-0153070P.
PR 23-SEP-1999; 99US-0153758P.
PR 24-SEP-1999; 99US-0154018P.
PR 28-SEP-1999; 99US-0154039P.
PR 29-SEP-1999; 99US-0154779P.
PR 04-OCT-1999; 99US-0155139P.
PR 05-OCT-1999; 99US-0155486P.
PR 06-OCT-1999; 99US-0155659P.
PR 07-OCT-1999; 99US-0156458P.
PR 08-OCT-1999; 99US-0156596P.
PR 12-OCT-1999; 99US-0157117P.
PR 13-OCT-1999; 99US-0157753P.
PR 13-OCT-1999; 99US-0157865P.
PR 13-OCT-1999; 99US-0158029P.
PR 14-OCT-1999; 99US-0158232P.
99US-0158369P.
99US-0159293P.
99US-0159294P.
99US-0159295P.
99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.:   2.81e-94   Length: 925
Score: 1355.50   Matches: 264
Percent Similarity: 96.00%   Conservative: 0
Best Local Similarity: 96.00%   Mismatches: 0
Query Match: 93.69%   Indels: 11
DB: 3   Gaps: 1

US-10-014-927-19 (1-279) x AAC49248 (1-925)

Qy   1   MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 134   ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG 193

Qy   21   CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle   40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 194   TGTTGAGGTTGAAGATCTCTTCTACAAGTATGGACCAATTGTGGACATTGATTTGAAGATT 253

Qy   41   ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgProAlaAspAsp   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 254   CCACCGAGACCTCCTGGTTATGCCTTTGTCGAGTTGAAGATCCTGTGATGCAGACGAT 313

Qy   61   AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla   80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 314   GCAATTTATGGACGTGATGGTTATGATTTTGATGGGTGTCGACTTCGGGTTGAAGATTCA 373

Qy   81   HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSer 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 374   CATGGTGGTCGTAGATTTTCACCATCAGTTGATAGGTACAGCAGCTACAGTGCCAGC 433

Qy 101   ArgAlaProSerArgArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 434   CGTGCACCTTCAAGACGCTCGACTACCGGCGTTCGTGTTGTGACCGGATTACCGCCTTCTGCT 493

Qy 121   SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGlyVal 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 494   TCGTGGCAGGACCTTAAGGATCACATGCGCAAAGCTGGAGATGTCTGCTTCTCTGAAGTT 553

Qy 141   PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 554   TTCCCTGACCGTAAAGGCATGTCTGGGGTTGTGGATTATAGCAACTATGATGATATGAAG 613

Qy 161   TyrAlaIleArgLysLeuAspAlaThrGluPheArgGlyAsnAlaPheSerArgAlaTyrIle 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 614   TACGCAATAAGGAAACTTGATGCCACTGAATTTCGAAATGCTTTCTCTAGTGCTTATATA 673

Qy 181   ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 674   CGGGTGAGGGAATATGAGTGAGTCGAAGCCCAGATGATTCTAAAAGCTAT 733

Qy 201   ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 734   AGAAGCAGGAGTCGGAGCCGTGGTCCAAGCTGTAGCTATAGTAGCAAGAGCAGGAGTGTG 793

Qy 221   SerProAlaArgSerIleSerProArgProLeuSerArgProLeuSerArgSerLeuTyr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 794   TCACCTGCTAGATCCATTTCCCGCGGTTCACGGCCCCTTAGTCGTTCTCGCTCGCTATAC 853

Qy 241   SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArg 260
          ||||||||||||||||||||||||||||||||||||||||||||--------------
Db 854   AGCTCTGTCTCAAGG---------------------------------TCCCAATCAAGA 880

Qy 261   SerLysSerArgSerArgSerArgSerAsnSerProValSerPro 275
          ||||||||||||||||||||||||||||||||||||||||||
Db 881   TCAAAATCAAGATCAAGATCAAGATCGAATTCTCCAGTTTCACCT 925

RESULT 2
AAC40180
ID   AAC40180 standard; DNA; 1213 BP.
XX
AC   AAC40180;
XX
DT   17-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 27352.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999; 99US-0121825P.
PR   05-MAR-1999; 99US-0123180P.
PR   09-MAR-1999; 99US-0123548P.
PR   23-MAR-1999; 99US-0125788P.
PR   25-MAR-1999; 99US-0126264P.
PR   29-MAR-1999; 99US-0126785P.
PR   01-APR-1999; 99US-0127462P.
PR   06-APR-1999; 99US-0128234P.
PR   08-APR-1999; 99US-0128714P.
PR   16-APR-1999; 99US-0129845P.
PR   19-APR-1999; 99US-0130077P.
PR   21-APR-1999; 99US-0130449P.
PR   23-APR-1999; 99US-0130510P.
PR   23-APR-1999; 99US-0130891P.
PR   28-APR-1999; 99US-0131449P.
PR   30-APR-1999; 99US-0132048P.
PR   30-APR-1999; 99US-0132407P.
PR   04-MAY-1999; 99US-0132484P.
PR   05-MAY-1999; 99US-0132485P.
PR   06-MAY-1999; 99US-0132486P.
PR   06-MAY-1999; 99US-0132487P.
PR   07-MAY-1999; 99US-0132863P.
PR   11-MAY-1999; 99US-0134256P.
PR   14-MAY-1999; 99US-0134218P.
PR   14-MAY-1999; 99US-0134219P.
PR   14-MAY-1999; 99US-0134221P.
PR   14-MAY-1999; 99US-0134370P.
PR   18-MAY-1999; 99US-0134768P.
PR   19-MAY-1999; 99US-0134941P.
PR   20-MAY-1999; 99US-0135153P.
PR   21-MAY-1999; 99US-0135353P.
PR   24-MAY-1999; 99US-0135629P.
PR   25-MAY-1999; 99US-0136021P.

```
PR  27-MAY-1999;   99US-0136392P.        PR  04-AUG-1999;   99US-0147204P.
PR  28-MAY-1999;   99US-0136782P.        PR  04-AUG-1999;   99US-0147302P.
PR  01-JUN-1999;   99US-0137222P.        PR  05-AUG-1999;   99US-0147192P.
PR  03-JUN-1999;   99US-0137528P.        PR  05-AUG-1999;   99US-0147260P.
PR  04-JUN-1999;   99US-0137502P.        PR  06-AUG-1999;   99US-0147303P.
PR  07-JUN-1999;   99US-0137724P.        PR  06-AUG-1999;   99US-0147416P.
PR  08-JUN-1999;   99US-0138094P.        PR  09-AUG-1999;   99US-0147493P.
PR  10-JUN-1999;   99US-0138540P.        PR  09-AUG-1999;   99US-0147495P.
PR  14-JUN-1999;   99US-0138847P.        PR  10-AUG-1999;   99US-0148171P.
PR  14-JUN-1999;   99US-0139119P.        PR  11-AUG-1999;   99US-0148319P.
PR  16-JUN-1999;   99US-0139452P.        PR  12-AUG-1999;   99US-0148341P.
PR  16-JUN-1999;   99US-0139453P.        PR  13-AUG-1999;   99US-0148585P.
PR  17-JUN-1999;   99US-0139492P.        PR  13-AUG-1999;   99US-0148684P.
PR  18-JUN-1999;   99US-0139454P.        PR  16-AUG-1999;   99US-0149368P.
PR  18-JUN-1999;   99US-0139455P.        PR  17-AUG-1999;   99US-0149175P.
PR  18-JUN-1999;   99US-0139456P.        PR  18-AUG-1999;   99US-0149426P.
PR  18-JUN-1999;   99US-0139457P.        PR  20-AUG-1999;   99US-0149722P.
PR  18-JUN-1999;   99US-0139458P.        PR  20-AUG-1999;   99US-0149723P.
PR  18-JUN-1999;   99US-0139459P.        PR  20-AUG-1999;   99US-0149929P.
PR  18-JUN-1999;   99US-0139460P.        PR  23-AUG-1999;   99US-0149902P.
PR  18-JUN-1999;   99US-0139461P.        PR  23-AUG-1999;   99US-0149930P.
PR  18-JUN-1999;   99US-0139462P.        PR  25-AUG-1999;   99US-0150566P.
PR  18-JUN-1999;   99US-0139463P.        PR  26-AUG-1999;   99US-0150884P.
PR  18-JUN-1999;   99US-0139750P.        PR  27-AUG-1999;   99US-0151065P.
PR  18-JUN-1999;   99US-0139763P.        PR  27-AUG-1999;   99US-0151066P.
PR  21-JUN-1999;   99US-0139817P.        PR  27-AUG-1999;   99US-0151080P.
PR  22-JUN-1999;   99US-0139899P.        PR  30-AUG-1999;   99US-0151303P.
PR  23-JUN-1999;   99US-0140353P.        PR  31-AUG-1999;   99US-0151438P.
PR  23-JUN-1999;   99US-0140354P.        PR  01-SEP-1999;   99US-0151930P.
PR  24-JUN-1999;   99US-0140695P.        PR  07-SEP-1999;   99US-0152363P.
PR  28-JUN-1999;   99US-0140823P.        PR  10-SEP-1999;   99US-0153070P.
PR  29-JUN-1999;   99US-0140991P.        PR  13-SEP-1999;   99US-0153758P.
PR  30-JUN-1999;   99US-0412877P.        PR  15-SEP-1999;   99US-0154018P.
PR  01-JUL-1999;   99US-0141842P.        PR  16-SEP-1999;   99US-0154039P.
PR  01-JUL-1999;   99US-0142154P.        PR  20-SEP-1999;   99US-0154779P.
PR  02-JUL-1999;   99US-0142055P.        PR  22-SEP-1999;   99US-0155139P.
PR  06-JUL-1999;   99US-0142803P.        PR  23-SEP-1999;   99US-0155486P.
PR  08-JUL-1999;   99US-0142920P.        PR  24-SEP-1999;   99US-0155659P.
PR  09-JUL-1999;   99US-0142977P.        PR  28-SEP-1999;   99US-0156458P.
PR  12-JUL-1999;   99US-0143542P.        PR  29-SEP-1999;   99US-0156596P.
PR  13-JUL-1999;   99US-0143624P.        PR  04-OCT-1999;   99US-0157117P.
PR  14-JUL-1999;   99US-0144005P.        PR  05-OCT-1999;   99US-0157753P.
PR  15-JUL-1999;   99US-0144085P.        PR  06-OCT-1999;   99US-0157865P.
PR  16-JUL-1999;   99US-0144086P.        PR  07-OCT-1999;   99US-0158029P.
PR  16-JUL-1999;   99US-0144325P.        PR  08-OCT-1999;   99US-0158222P.
PR  19-JUL-1999;   99US-0144331P.        PR  12-OCT-1999;   99US-0158369P.
PR  19-JUL-1999;   99US-0144333P.        PR  13-OCT-1999;   99US-0159293P.
PR  19-JUL-1999;   99US-0144334P.        PR  13-OCT-1999;   99US-0159294P.
PR  19-JUL-1999;   99US-0144335P.        PR  13-OCT-1999;   99US-0159295P.
PR  20-JUL-1999;   99US-0144352P.        PR  14-OCT-1999;   99US-0159329P.
PR  20-JUL-1999;   99US-0144632P.        PR  14-OCT-1999;   99US-0159330P.
PR  21-JUL-1999;   99US-0144884P.        PR  14-OCT-1999;   99US-0159331P.
PR  21-JUL-1999;   99US-0144814P.        PR  18-OCT-1999;   99US-0159584P.
PR  21-JUL-1999;   99US-0450865P.        PR  21-OCT-1999;   99US-0160741P.
PR  21-JUL-1999;   99US-0450866P.        PR  21-OCT-1999;   99US-0160767P.
PR  21-JUL-1999;   99US-0450888P.        PR  21-OCT-1999;   99US-0160768P.
PR  22-JUL-1999;   99US-0450887P.        PR  21-OCT-1999;   99US-0160769P.
PR  22-JUL-1999;   99US-0450899P.        PR  21-OCT-1999;   99US-0160770P.
PR  22-JUL-1999;   99US-0451192P.        PR  21-OCT-1999;   99US-0160814P.
PR  23-JUL-1999;   99US-0451145P.        PR  21-OCT-1999;   99US-0160815P.
PR  23-JUL-1999;   99US-0452218P.        PR  22-OCT-1999;   99US-0160980P.
PR  23-JUL-1999;   99US-0452224P.        PR  22-OCT-1999;   99US-0160981P.
PR  26-JUL-1999;   99US-0452276P.        PR  22-OCT-1999;   99US-0160989P.
PR  27-JUL-1999;   99US-0459139P.        PR  25-OCT-1999;   99US-0161404P.
PR  27-JUL-1999;   99US-0459189P.        PR  25-OCT-1999;   99US-0161405P.
PR  27-JUL-1999;   99US-0459191P.        PR  25-OCT-1999;   99US-0161406P.
PR  28-JUL-1999;   99US-0459518P.        PR  26-OCT-1999;   99US-0161359P.
PR  02-AUG-1999;   99US-0146386P.        PR  26-OCT-1999;   99US-0161360P.
PR  02-AUG-1999;   99US-0146389P.        PR  26-OCT-1999;   99US-0161361P.
PR  03-AUG-1999;   99US-0147038P.        PR  28-OCT-1999;   99US-0161920P.
                                         PR  28-OCT-1999;   99US-0161992P.
                                         PR  28-OCT-1999;   99US-0161993P.
```

PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:   1213
Score: 1359.50   Matches: 267
Percent Similarity: 95.02%   Conservative: 0
Best Local Similarity: 95.02%   Mismatches: 1
Query Match: 93.69%   Indels: 13
DB: Gaps: 2

US-10-014-927-19 (1-279) x AAC40180 (1-1213)

```
Qy      1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys    20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGGCTGGAGATATTCGCAAG   193

Qy     21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle    40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 TGTGAGGTTGAAGATCTCTTCTACAAGTATGGACCAATTGTGGACATTGATTTGAAGATT   253

Qy     41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProAspAlaAspAspAsp    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 CCACCGCGACCTCCTGGTTATGCCTTTGTCGAGTTGGAAGATCCTGATGCAGACGAT    313

Qy     61 AlaIleTyrGlyArgAspGlyTyrArgAspPheAspGlyCysArgLeuArgValGluIleAla    80
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 GCAATTTATGGACGTGATGGTTATGATTTTGATGGGTGTCGACTTCGGGTTGAGATTGCA   373

Qy     81 HisGlyGlyArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer    100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 CATGGTGGTCGTAGATTTTCACCATCAGTTGATAGGTACAGCAGTCACAGTGCGAGC   433

Qy    101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 CGTGCACCTTCAAGACGCTCTGACTATCGCGTGCTTGTGACCGGATTACCGCCTTCTGCT   493

Qy    121 SerTrpGlnAsnAspLeuLysAspHisMetArgLysAlaGlyValCysPheSerGluVal   140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    494 TCGTGGCAGAACGACCTTAAGGATCACATGCGCAAAGCTGGAGATGTCTGCTTCTCTGAAGTT   553

Qy    141 PheProAspLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys   160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    554 TTCCCTGACCGTAAAGGCATGTCTGGGGTTGTGGATTATAGCAACTATGATGATGAAG   613

Qy    161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    614 TACGCAATAAGGAAACTTGATGCGACCGAATTTCGAAATGCTTTCTCTAGTGCTTATATA   673

Qy    181 ArgValArgGluTyrGluSerArgValSerArgSerValSerProAspAspSerLysSerTyr   200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    674 CGGGTGAGGGAATATGAGTCGAGGGAGTGTGAGTCCAGATGATTCTAAAAGCTAT   733

Qy    201 ArgSerArgSerArgSerArgSerGlyTyrProSerCysSerTyrSerSerLysSerSerVal   220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    734 AGAAGCAGGAGTCGGAGCCGTGGTCCAAGTGTGCTATATAGCAAGAACAGCAGGAGTGTG   793

Qy    221 SerProAlaArgSerIleSerArgArgProLeuSerArgSerArgSerArgLeuTyr   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    794 TCACCTGCTAGATCCATTTCCCGCGGTTCACGGCCCCTTAGTCGTTCCGCTCGCCATAC   853

Qy    241 SerSerValSerArgSerArgGlySerLeuLeuArgAlaGlyAlaGlyAspTrpIleSerGlnSerArg   260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    854 AGCTCTGTCTCAAGG--------------------------TCCCAATCAAGA   880

Qy    261 SerLysSerArgSer-------ArgSerArgSerAsnSerProValSerProValIleSer   278
          ||||||||||||||                ||||||||||||||||||||||||||
Db    881 TCAAAATCAAGATCAACAACAYCACGATCAAGATCGAATTCTCCAGTTTCACCTGTGATATCT   940

Qy    279 Gly   279
          |||
Db    941 GGA   943
```

RESULT 3

AAC49242
ID AAC49242 standard; DNA; 771 BP.
XX
AC AAC49242;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60449.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.

PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
PR   16-JUL-1999;   99US-0144086P.
PR   19-JUL-1999;   99US-0144325P.
PR   19-JUL-1999;   99US-0144331P.
PR   19-JUL-1999;   99US-0144332P.
PR   19-JUL-1999;   99US-0144333P.
PR   19-JUL-1999;   99US-0144334P.
PR   19-JUL-1999;   99US-0144335P.
PR   20-JUL-1999;   99US-0144352P.
PR   20-JUL-1999;   99US-0144632P.
PR   21-JUL-1999;   99US-0144484P.
PR   21-JUL-1999;   99US-0144814P.
PR   21-JUL-1999;   99US-0145086P.
PR   21-JUL-1999;   99US-0145088P.
PR   22-JUL-1999;   99US-0145085P.
PR   22-JUL-1999;   99US-0145087P.
PR   22-JUL-1999;   99US-0145089P.
PR   22-JUL-1999;   99US-0145192P.
PR   23-JUL-1999;   99US-0145145P.
PR   23-JUL-1999;   99US-0145218P.
PR   23-JUL-1999;   99US-0145224P.
PR   26-JUL-1999;   99US-0145276P.
PR   27-JUL-1999;   99US-0145913P.
PR   27-JUL-1999;   99US-0145918P.
PR   27-JUL-1999;   99US-0145919P.
PR   28-JUL-1999;   99US-0145951P.
PR   02-AUG-1999;   99US-0146386P.
PR   02-AUG-1999;   99US-0146388P.
PR   02-AUG-1999;   99US-0146389P.
PR   03-AUG-1999;   99US-0147038P.
PR   04-AUG-1999;   99US-0147204P.
PR   04-AUG-1999;   99US-0147302P.
PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.

PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150556P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   30-AUG-1999;   99US-0151080P.
PR   31-AUG-1999;   99US-0151303P.
PR   01-SEP-1999;   99US-0151438P.
PR   07-SEP-1999;   99US-0151930P.
PR   10-SEP-1999;   99US-0152363P.
PR   13-SEP-1999;   99US-0153070P.
PR   15-SEP-1999;   99US-0153758P.
PR   16-SEP-1999;   99US-0154018P.
PR   20-SEP-1999;   99US-0154039P.
PR   22-SEP-1999;   99US-0154779P.
PR   23-SEP-1999;   99US-0155139P.
PR   24-SEP-1999;   99US-0155486P.
PR   28-SEP-1999;   99US-0155659P.
PR   29-SEP-1999;   99US-0156458P.
PR   04-OCT-1999;   99US-0156596P.
PR   05-OCT-1999;   99US-0157117P.
PR   06-OCT-1999;   99US-0157753P.
PR   07-OCT-1999;   99US-0157865P.
PR   08-OCT-1999;   99US-0158029P.
PR   12-OCT-1999;   99US-0158232P.
PR   13-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   14-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   18-OCT-1999;   99US-0159637P.
PR   21-OCT-1999;   99US-0159638P.
PR   21-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   22-OCT-1999;   99US-0160814P.
PR   22-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   25-OCT-1999;   99US-0160981P.
PR   25-OCT-1999;   99US-0160989P.
PR   26-OCT-1999;   99US-0161404P.
PR   26-OCT-1999;   99US-0161405P.
PR   28-OCT-1999;   99US-0161406P.
PR   28-OCT-1999;   99US-0161359P.
PR   29-OCT-1999;   99US-0161360P.
PR   29-OCT-1999;   99US-0161361P.
PR   29-OCT-1999;   99US-0161920P.
PR   29-OCT-1999;   99US-0161992P.
PR   29-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

Alignment Scores:
Pred. No.:   4.06e-93
Score: 1343.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 92.56%
DB: 3

Length: 771
Matches: 256
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-014-927-19 (1-279) x AAC49242 (1-771)

Qy   1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCTGGGAGATATTCGCAAG   60

Qy 21 CysGluValGluAspPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle   40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 TGTGAGGTTGAAGATCTTCTTCTTCAAGTATGTGGACATTGATTTGAAGATT   120

```
XX
DR   WPI; 2002-304127/34.
XX
PT   Identifying a stress condition to which a plant cell has been exposed and
PT   producing plants with increased tolerance to these abiotic stresses.
XX
PS   Claim 144; SEQ ID NO 1199; 577pp + Sequence Listing; English.
XX
CC   The invention relates to identifying a stress condition to which a plant
CC   cell has been exposed, comprising: (a) contacting nucleic acid
CC   representative of expressed polynucleotides in the plant cell with an
CC   array or probes representative of the plant cell genome; and (b)
CC   detecting a profile of expressed polynucleotides in the plant cell
CC   characteristic of a stress response. The method is useful in the
CC   production of transgenic plants, cells and seeds and in producing plants
CC   with increased tolerance to abiotic stress. The present sequence is that
CC   of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC   in methods of the invention. Note: The sequence data for this patent is
CC   not represented in the printed specification but is based on sequence
CC   information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 762 BP; 189 A; 161 C; 197 G; 215 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score:                    9.84e-88          Length:            762
                          1272.00           Matches:           253
Percent Similarity:       91.67%            Conservative:      0
Best Local Similarity:    91.67%            Mismatches:        1
Query Match:              87.66%            Indels:            23
DB:                                         Gaps:              2

US-10-014-927-19 (1-279) x ABZ13394 (1-762)

Qy    1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   20
Db    1 ATGAGTAGCCGATGGAACGTACGAATCATAGTTGTGGGAATTTGCCTGGAGATATTCGCAAG   60

Qy   21 CysGluValGluAspLeuProPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle   40
Db   61 TGTGAGGTTGAAGATCTCCCATTCTACAAGTATGGACCAATTGTGGACATTGATTTGAAGATT  120

Qy   41 ProProProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp   60
Db  121 CCACCGAGACCTCCTGGTTATGCCTTTGTCGAGTTTGAAGATCCCCGTGATGCAGACGAT  180

Qy   61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla   80
Db  181 GCAATTTATGGACGTGATGGTTATGATTTTGATGGGTGTCGGGTTGCCTGGAGATATTGCA  240

Qy   81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer  100
Db  241 CATGGTGGTCGTAGATTTCACCATCAGTTGATAGGTACAGCAGCTACAGTGCCGAGC  300

Qy  101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla  120
Db  301 CGTGCACCTTCAAGACGCTCTGACTACCGCGTGCTTGTGACCGGATTACCGCCTTCTGCT  360

Qy  121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal  140
Db  361 TCGTGGCAGGACCTTAAGGATCACATGCGCAAAGCTGGAGATGTGTGCTTCTCTGAAGTT  420

Qy  141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys  160
Db  421 TTCCCTGACCGTAAAGGCATGTCTGGGGTTGTGGATTATAGCAACTATGATGATATGAAG  480

Qy  161 TyrAlaIleArgLysLeuAspAlaThrGluHisArgSerArgSerAlaPheSerArgTyrIle  180
Db  481 TACGCAATAAGGAAACTTGATGCCACTGAATTTCGAAATGCTTTCTCTAGTGCTTATATA  540

Qy  181 ArgValArgGluSerArgGluSerValSerArgSerProAspAspSerLysSerSerTyr  200
Db  541 CGGGTGAGGGAATATGAGTCAGAGAGTGTGAGTCGAAGCCCAGATGATTCTAAAGCTAT  600

Qy  201 ArgSerArgSerArgSerArgGlyProSerArgTyrSerArgCysSerSerArgSerVal  220
Db  601 AGAAGCAGGAGCAGGAGCCGTGGTCCAAGCTGTTAGCTATAGTAGCAAGAGCAGGAGTGTG  660

Qy  221 SerProAlaArgSerIleSerProArgSerProLeuSerArgSerArgSerArgSerLeuTyr  240
Db  661 TCACCTGCTAGATCCATTTCCCCGCGTTCACCGCTCCCTTAGTCGTTCCGCTCGCTATAC  720

Qy  241 SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAlaGlyAspTrpIle  256
Db  721 AGCTCTGTCTCAAGATCTCGGCTCACTGCTACGAGCGCTGGGGATTGGATC  768

RESULT 4
AB213394
ID   AB213394 standard; DNA; 762 BP.
XX
AC   AB213394;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 1199.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US026685.
XX
PR   24-AUG-2000; 2000US-0227866P.
PR   26-JAN-2001; 2001US-0264647P.
PR   22-JUN-2001; 2001US-0300111P.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Harper JF, Kreps J, Wang X, Zhu T;
```

```
Qy   201 ArgSerArgGlyProSerCysSerTyrSerSerLysSerArgVal 220
             |||||||||||                              |||
Db   601 AGAAGCAGGAGTCGG----------------------AGCCGTGT-GTG 626

Qy   221 SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyr 240
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   627 TCACCTGCTAGATCCATTTCCCGCGTTCACGGCCCCTTAGTCGTTCTCGCTCGCTATAC 686

Qy   241 SerSerValSerArgSerGlySerLeuArgAlaGlyAspTrpIleSerGlnSerArg 260
             |||||||||||||||||                          |||||||||||
Db   687 AGCTCTGTCTCAAGG----------------------TCCCAATCAAGA 713

Qy   261 SerLysSerSerArgSerArgSerArgSerAsnSerProValSerProVal 276
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   714 TCAAAATCAAGATCAAGATCAAGATCGAATTCTCCAGTTTCACCTGTA 761

RESULT 5
AAC42865
ID   AAC42865 standard; DNA; 714 BP.
XX
AC   AAC42865;
XX
DT   17-OCT-2000   (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 37144.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;    99US-0121825P.
PR   05-MAR-1999;    99US-0123180P.
PR   09-MAR-1999;    99US-0123548P.
PR   23-MAR-1999;    99US-0125788P.
PR   25-MAR-1999;    99US-0126264P.
PR   29-MAR-1999;    99US-0126785P.
PR   01-APR-1999;    99US-0127462P.
PR   06-APR-1999;    99US-0128234P.
PR   08-APR-1999;    99US-0128714P.
PR   16-APR-1999;    99US-0129845P.
PR   19-APR-1999;    99US-0130077P.
PR   21-APR-1999;    99US-0130449P.
PR   23-APR-1999;    99US-0130510P.
PR   23-APR-1999;    99US-0130891P.
PR   28-APR-1999;    99US-0131449P.
PR   30-APR-1999;    99US-0132048P.
PR   30-APR-1999;    99US-0132407P.
PR   04-MAY-1999;    99US-0132484P.
PR   05-MAY-1999;    99US-0132485P.
PR   06-MAY-1999;    99US-0132486P.
PR   06-MAY-1999;    99US-0132487P.
PR   07-MAY-1999;    99US-0132863P.
PR   11-MAY-1999;    99US-0134256P.
PR   14-MAY-1999;    99US-0134218P.
PR   14-MAY-1999;    99US-0134219P.
PR   14-MAY-1999;    99US-0134221P.
PR   14-MAY-1999;    99US-0134370P.
PR   18-MAY-1999;    99US-0134768P.
PR   19-MAY-1999;    99US-0134941P.
PR   20-MAY-1999;    99US-0135124P.
PR   21-MAY-1999;    99US-0135353P.
PR   24-MAY-1999;    99US-0135629P.
PR   25-MAY-1999;    99US-0136021P.
PR   27-MAY-1999;    99US-0136392P.
PR   28-MAY-1999;    99US-0136782P.
PR   01-JUN-1999;    99US-0137222P.
PR   03-JUN-1999;    99US-0137528P.
PR   04-JUN-1999;    99US-0137502P.
PR   07-JUN-1999;    99US-0137724P.
PR   08-JUN-1999;    99US-0138094P.
PR   10-JUN-1999;    99US-0138540P.
PR   10-JUN-1999;    99US-0138847P.
PR   14-JUN-1999;    99US-0139119P.
PR   16-JUN-1999;    99US-0139452P.
PR   16-JUN-1999;    99US-0139453P.
PR   17-JUN-1999;    99US-0139492P.
PR   18-JUN-1999;    99US-0139454P.
PR   18-JUN-1999;    99US-0139455P.
PR   18-JUN-1999;    99US-0139456P.
PR   18-JUN-1999;    99US-0139457P.
PR   18-JUN-1999;    99US-0139458P.
PR   18-JUN-1999;    99US-0139459P.
PR   18-JUN-1999;    99US-0139460P.
PR   18-JUN-1999;    99US-0139461P.
PR   18-JUN-1999;    99US-0139462P.
PR   18-JUN-1999;    99US-0139463P.
PR   18-JUN-1999;    99US-0139750P.
PR   18-JUN-1999;    99US-0139763P.
PR   21-JUN-1999;    99US-0139817P.
PR   22-JUN-1999;    99US-0139899P.
PR   23-JUN-1999;    99US-0140353P.
PR   23-JUN-1999;    99US-0140354P.
PR   24-JUN-1999;    99US-0140695P.
PR   28-JUN-1999;    99US-0140823P.
PR   29-JUN-1999;    99US-0140991P.
PR   30-JUN-1999;    99US-0141287P.
PR   01-JUL-1999;    99US-0141842P.
PR   01-JUL-1999;    99US-0142154P.
PR   02-JUL-1999;    99US-0142055P.
PR   06-JUL-1999;    99US-0142390P.
PR   08-JUL-1999;    99US-0142803P.
PR   09-JUL-1999;    99US-0142920P.
PR   12-JUL-1999;    99US-0142977P.
PR   13-JUL-1999;    99US-0143542P.
PR   14-JUL-1999;    99US-0143624P.
PR   15-JUL-1999;    99US-0144005P.
PR   16-JUL-1999;    99US-0144085P.
PR   16-JUL-1999;    99US-0144086P.
PR   19-JUL-1999;    99US-0144325P.
PR   19-JUL-1999;    99US-0144331P.
PR   19-JUL-1999;    99US-0144332P.
PR   19-JUL-1999;    99US-0144333P.
PR   19-JUL-1999;    99US-0144334P.
PR   19-JUL-1999;    99US-0144335P.
PR   20-JUL-1999;    99US-0144352P.
PR   20-JUL-1999;    99US-0144632P.
PR   20-JUL-1999;    99US-0144484P.
PR   21-JUL-1999;    99US-0144814P.
PR   21-JUL-1999;    99US-0145086P.
PR   21-JUL-1999;    99US-0144332P.
PR   22-JUL-1999;    99US-0145085P.
PR   22-JUL-1999;    99US-0145087P.
PR   22-JUL-1999;    99US-0145088P.
PR   22-JUL-1999;    99US-0145089P.
PR   22-JUL-1999;    99US-0145192P.
PR   23-JUL-1999;    99US-0145145P.
PR   23-JUL-1999;    99US-0145218P.
PR   23-JUL-1999;    99US-0145224P.
PR   23-JUL-1999;    99US-0145276P.
PR   26-JUL-1999;    99US-0145913P.
PR   27-JUL-1999;    99US-0145918P.
PR   27-JUL-1999;    99US-0145919P.
PR   28-JUL-1999;    99US-0145951P.
PR   02-AUG-1999;    99US-0146386P.
PR   02-AUG-1999;    99US-0146388P.
PR   03-AUG-1999;    99US-0146389P.
PR   04-AUG-1999;    99US-0147038P.
PR   04-AUG-1999;    99US-0147204P.
PR   04-AUG-1999;    99US-0147302P.
```

Alignment Scores:
Pred. No.:　　　　　　　　　　3.46e-79
Score:　　　　　　　　　　　　1159.00
Percent Similarity:　　　　100.00%
Best Local Similarity:　　100.00%
Query Match:　　　　　　　　79.88%
DB:

Length:　　　　714
Matches:　　　218
Conservative:　　0
Mismatches:　　　0
Indels:　　　　　0
Gaps:　　　　　　0

US-10-014-927-19 (1-279) x AAC42865 (1-714)

```
Qy    1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG   60

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle   40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TGTGAGGTTGAAGATCTCTTCTACAAGTATGGACCAATTGTGGACATTGATTTGAAGATT  120

Qy   41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp   60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CCACCGAGACCCTCCGGTTATGCCTTTGTCGAGTTTGAAGATCCTCGTGATGCAGACGAT  180

Qy   61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla   80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GCAATTTATGGACGTGATGGTTATGATTTTGATGGGTGTCGACTTCGGGTTGAGATTGCA  240

Qy   81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer    100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CATGGTGGTCGTAGATTTTCACCATCAGTTGATAGGTACAGCAGCACTACAGTGCCGAGC  300

Qy  101 ArgAlaProSerArgArgAspTyrArgValLeuValThrGlyLeuProProSerAla    120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 CGTGCACCTTCAAGACGCTCTGACTACCGGTGCTTGTTGACCGGATTACCGCCTTCTGCT  360

Qy  121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal  140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 TCGTGGCAGGACCTTAAGGATCACATGCGCAAAGCTGGAGATGTCTGCTTCTCTGAAGTT  420

Qy  141 PheProAspArgLysGlyMetSerGlyValValValAspTyrSerAsnTyrArgAspMetLys  160
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 TTCCCTGACCGTAAAGGCATGTCTGGGGTTGTGGATTATAGCAACTATGATGATATGAAG  480

Qy  161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerArgAlaTyrIle  180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 TACGCCAATAAGGAAAACTTGATGCCCACTGAATTTCGAAATGCTTTCTCTAGTGCTTATATA  540

Qy  181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr  200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CGGGTGAGGGAATATGAGTCGAGGAGTGTGAGTCGAAGCCCAGATGATTCTAAAAGCTAT  600

Qy  201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerSerArg    218
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 AGAAGCAGGAGTCGGAGCCGTGGTCCAAGCTGTAGCTATAGTAGCAAGAGACAGG  654
```

RESULT 6
AAC81899
ID   AAC81899　standard; DNA; 4044 BP.
XX
AC   AAC81899;
XX
DT   23-FEB-2001　(first entry)
XX
DE   A. thaliana SRp30 cDNA.
XX
KW   SR protein; splice-factor activity; plant; developmental behavior;
KW   flowering; crop plant; cereal; bean; rice; fruit; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200065059-A1.
XX
PD   02-NOV-2000.
XX

PR   05-AUG-1999;   99US-0147192P.
PR   05-AUG-1999;   99US-0147260P.
PR   06-AUG-1999;   99US-0147303P.
PR   06-AUG-1999;   99US-0147416P.
PR   09-AUG-1999;   99US-0147493P.
PR   09-AUG-1999;   99US-0147935P.
PR   10-AUG-1999;   99US-0148171P.
PR   11-AUG-1999;   99US-0148319P.
PR   12-AUG-1999;   99US-0148341P.
PR   13-AUG-1999;   99US-0148565P.
PR   13-AUG-1999;   99US-0148684P.
PR   16-AUG-1999;   99US-0149368P.
PR   17-AUG-1999;   99US-0149175P.
PR   18-AUG-1999;   99US-0149426P.
PR   20-AUG-1999;   99US-0149722P.
PR   20-AUG-1999;   99US-0149723P.
PR   20-AUG-1999;   99US-0149929P.
PR   23-AUG-1999;   99US-0149902P.
PR   23-AUG-1999;   99US-0149930P.
PR   25-AUG-1999;   99US-0150566P.
PR   26-AUG-1999;   99US-0150884P.
PR   27-AUG-1999;   99US-0151065P.
PR   27-AUG-1999;   99US-0151066P.
PR   27-AUG-1999;   99US-0151080P.
PR   30-AUG-1999;   99US-0151303P.
PR   31-AUG-1999;   99US-0151438P.
PR   01-SEP-1999;   99US-0151930P.
PR   07-SEP-1999;   99US-0153363P.
PR   10-SEP-1999;   99US-0153070P.
PR   15-SEP-1999;   99US-0153758P.
PR   15-SEP-1999;   99US-0154018P.
PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0155596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   13-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

```
PF   20-APR-2000; 2000WO-AT000100.
XX
PR   23-APR-1999; 99AT-00000727.
XX
PA   (OSTP ) OESTERR FORSCH SEIBERSDORF.
XX
PI   Barta A, Lopato S, Kalyna M, Dorner S;
XX
DR   WPI; 2000-687349/67.
DR   P-PSDB; AAB11414, AAB11419.
XX
PT   Novel proteins with splice-factor activity in plants, useful e.g. for
PT   altering flowering time or development, and the nucleic acid that encodes
PT   it.
XX
PS   Claim 5; Fig 1A; 67pp; German.
XX
CC   This invention describes a novel protein (I) with splice-factor activity
CC   in plants (I) modifies the choice of splice sites in many plant pre-
CC   mRNAs. (I) (also the nucleic acid that encodes them and related vectors
CC   or expression systems) are used: (i) to alter splice patterns in plants,
CC   or their parts; (ii) to alter developmental behavior of plants; and/or
CC   (iii) to delay flowering, particularly by at least 25% relative to the
CC   wild type, especially in crop plants such as cereals, beans, rice and
CC   fruit
XX
SQ   Sequence 4044 BP; 1061 A; 728 C; 828 G; 1427 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score:                 9.82e-57
                       875.00
Percent Similarity:    43.53%          Length:        4044
Best Local Similarity: 43.53%          Matches:        242
Query Match:           60.30%          Conservative:     0
DB:                    3               Mismatches:       3
                                       Indels:         314
                                       Gaps:             9

US-10-014-927-19 (1-279) x AAC81899 (1-4044)

Qy    1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db  808 ATGAGTAGCCGATGGAATCGTACGATCTATATTGTGGGAAATTTGCCTGGAGATATTCGCAAG 867

Qy   21 CysGlnValGluSerAspLeuPheTyrLys--------------------- 29
Db  868 TGTGAGGTTGAAGATCTCTTCTACAGGTTTGAAAATTTCCTCTTTTCTCCTCGATAAAAA 927

Qy   29 ------------------------- 29
Db  928 TTGAATTCATTATGACTAGTTTGGGTTCATAAATTCGTCTCGTGCTGAGACAAT 987

Qy   30 -----------------TyrGlyProIleValAspIleAspLeuLy 39
Db  988 TTAAATCGACTCTTATGTATATTGTTCAGTATGGGACAATTGATTGAA 1047

Qy   39 sIleProProArgProProGlyTyrAlaPheValGlu------------- 51
Db 1048 GATTCCACCGAGACCTCCGGTTATGCCTTTGTCGA-GGTATATTGATCAAGTACAAATT 1106

Qy   51 --------------------------------------- 51
Db 1107 TGTTTTTTTCTTCTTGTAATAGTAATAGTAATGACTAAGATAGTTTGTTATTGGTG 1166

Qy   52 ---PheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyrAspPh 70
Db 1167 GCAGTTTGAAGATCCTCGATGCAGACGATGCAATTTATGGAACGTGATGGTTATGGATTT 1226

Qy   70 eAspGlyCysArgLeuArg------------------- 76
Db 1227 TGATGGGTGTCGACTTTCGGGTTAGTAACGCATGATGAAAGCTAGTTTTCTGTAA 1286

Qy   77 -------------ValGluIleAlaHisGlyGlyAla 84
Db 1287 TTTCTTGTAAGGTGTATCTTTGTGTGATGTTTTAGGTTGAGATTGCAATGGTGGTC 1346

Qy   84 rgArgPheSerProValAspArgTyrSerSerSerTyrSerAlaSerArgAlaProS 104
Db 1347 GTAGATTTTCACCATCAGTTGATAGGTACAGCAGCTACAGCAGCCGTGCACCTT 1406

Qy  104 erArgArgSerAspTyrArg 110
Db 1407 CAAGACGCCTCTGACTACCGCGGTTGTAGAGTCTTCTGATTGTGTTATTGGTGTG 1466

Qy  111 --------------------ValLeuValThrGly 115
Db 1467 TAAAATTTTATATTGAAAACTCATTTTACTACCTAAACATGTAGTGCTTGGACCGGA 1526

Qy  116 LeuProProSerAlaSerTrpGlnAspLeuLys-------------- 126
Db 1527 TTACCGCCCTTCTGCTTCGTGGCAGGACCTTAA-GGTAAGGGACACTATATAGTCTTTTTC 1585

Qy  126 ------------------- 126
Db 1586 TCTGAATGTTGGTTCTCTATATCATGTTTGGATTTATCTCTTTTCTGAAAATGATGTTAT 1645

Qy  127 ---AspHisMetArgLysAlaGlyAspValCysPheSerGlu 139
Db 1646 TTGCTATTTACGGGTGATTAGGATCACATGCCAAAGCTGGAGATGTCTGCTTCTCTGAA 1705

Qy  140 ValPheProAspArg-Lys 145
Db 1706 GTTTTCCCTGACCCTAAAGGTGAGTTGACATTCGATAGTTGGATAAGCTTTTTGATTGA 1765

Qy  145 ------------------- 145
Db 1766 TGTGTTAGTAAATTAGTCTTTGTGAAGAGAATAGGTGTTAAGCATCTGAACTGCTAAAC 1825

Qy  146 ---GlyMetSerGlyValValAspTyrSerAsnTyrArgAspAs 158
Db 1826 TCACATTCAGTATTCTTTGTAGCGATGTCTCTGGGGTTGTGTGGATTATAGCAACTATGATGA 1885

Qy  158 pMetLysTyrAla- 162
Db 1886 TATGAAGTACGGC-AGTAAGTTTTATATCTTGCAACGCAAATGTTCCTGGACTTATGCCT 1944

Qy  162 ------------------- 162
Db 1945 TAGACTGCTTTGTTTGTTCATAGTATACCGAGTCTGAATTATCTTCCTGGAGGCCAGTGTTG 2004

Qy  163 ---IleArgLysLeuAspAlaThrGl 170
Db 2005 GATCTTTGATGTGTCCCTTAAAATTTTGATGGTGACAGATAAGGAAACTTGATGCCACTGA 2064

Qy  170 uPheArgAsnAlaPheSerSerAlaTyrIleArg--- 181
Db 2065 ATTTCGAAATGCTTTTCCTCCAGTGCTTATATACGCGTATGTTGTATTGCTTCTTTGATTT 2124

Qy  182 ------------------ValArgGluTyrG 186
Db 2125 TGTTAAGCATAAGTGGATATGGAGTCATCTCGAATTTACTGTTCAGGTGGAGGGAATATG 2184

Qy  186 luSerArgValSerArgCysSerLysSerArgSerTyrArgSerArgSerArg 206
Db 2185 AGTCGAGGAGTGTGAGTCAAGCCCAGATGATTCTAAAAGCTATAGAAGCAGGAGTCCGGA 2244

Qy  206 erArgGlyProSerCysSerCysSerSerLysSerArg------------ 218
Db 2245 GCCGTGGTCCAAGCTGTAGCTATAGTAGCAAGACGAGCAGGAGGTCTCTTTTTTTTTTTTT 2304

Qy  218 ---------- 218
Db 2305 TTCATAAACCTAAGACATATAAGGGATTTTTATTGTAACTTATTTATGAAATTAACTGAC 2364

Qy  219 ---SerValSerProAlaArgSerIleSerProArgSerArgProLeu 233
Db 2365 TTCTAAATGCAATGCAGTGTGTCACCCGTCTAGATCCATTTCCCGCGTTCACGGCGCCCTT 2424
```

```
Qy    234 SerArgSerArgSerLeuTyrSerSerValSerArg 245
          |||||||||||||||||||||||||||||||||||
Db   2425 AGTCGTTCTCGCTCGCTATACAGCTCTGTCTCAAGG 2460

RESULT 7
AAC34814
ID   AAC34814 standard; DNA; 1232 BP.
XX
AC   AAC34814;
XX
DT   17-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 7984.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999; 99US-0121825P.     PR   18-JUN-1999; 99US-0139455P.
PR   05-MAR-1999; 99US-0123180P.     PR   18-JUN-1999; 99US-0139456P.
PR   09-MAR-1999; 99US-0123548P.     PR   18-JUN-1999; 99US-0139457P.
PR   23-MAR-1999; 99US-0125788P.     PR   18-JUN-1999; 99US-0139458P.
PR   25-MAR-1999; 99US-0126264P.     PR   18-JUN-1999; 99US-0139459P.
PR   29-MAR-1999; 99US-0126785P.     PR   18-JUN-1999; 99US-0139460P.
PR   01-APR-1999; 99US-0127462P.     PR   18-JUN-1999; 99US-0139461P.
PR   06-APR-1999; 99US-0128234P.     PR   18-JUN-1999; 99US-0139462P.
PR   08-APR-1999; 99US-0128714P.     PR   18-JUN-1999; 99US-0139463P.
PR   16-APR-1999; 99US-0129845P.     PR   18-JUN-1999; 99US-0139750P.
PR   19-APR-1999; 99US-0130077P.     PR   18-JUN-1999; 99US-0139763P.
PR   21-APR-1999; 99US-0130449P.     PR   21-JUN-1999; 99US-0139817P.
PR   23-APR-1999; 99US-0130510P.     PR   22-JUN-1999; 99US-0139899P.
PR   23-APR-1999; 99US-0130891P.     PR   23-JUN-1999; 99US-0140353P.
PR   28-APR-1999; 99US-0131449P.     PR   23-JUN-1999; 99US-0140354P.
PR   30-APR-1999; 99US-0132048P.     PR   24-JUN-1999; 99US-0140695P.
PR   30-APR-1999; 99US-0132407P.     PR   28-JUN-1999; 99US-0140823P.
PR   04-MAY-1999; 99US-0124484P.     PR   29-JUN-1999; 99US-0140991P.
PR   05-MAY-1999; 99US-0132485P.     PR   30-JUN-1999; 99US-0141287P.
PR   06-MAY-1999; 99US-0132486P.     PR   01-JUL-1999; 99US-0141842P.
PR   06-MAY-1999; 99US-0132487P.     PR   01-JUL-1999; 99US-0142154P.
PR   07-MAY-1999; 99US-0132863P.     PR   02-JUL-1999; 99US-0142055P.
PR   11-MAY-1999; 99US-0134256P.     PR   06-JUL-1999; 99US-0142390P.
PR   14-MAY-1999; 99US-0134218P.     PR   08-JUL-1999; 99US-0142803P.
PR   14-MAY-1999; 99US-0134219P.     PR   09-JUL-1999; 99US-0142920P.
PR   14-MAY-1999; 99US-0134221P.     PR   12-JUL-1999; 99US-0142977P.
PR   14-MAY-1999; 99US-0134370P.     PR   13-JUL-1999; 99US-0143542P.
PR   18-MAY-1999; 99US-0134768P.     PR   14-JUL-1999; 99US-0143624P.
PR   19-MAY-1999; 99US-0134941P.     PR   15-JUL-1999; 99US-0144005P.
PR   20-MAY-1999; 99US-0135124P.     PR   16-JUL-1999; 99US-0144085P.
PR   21-MAY-1999; 99US-0135353P.     PR   16-JUL-1999; 99US-0144086P.
PR   24-MAY-1999; 99US-0135629P.     PR   19-JUL-1999; 99US-0144325P.
PR   25-MAY-1999; 99US-0136021P.     PR   19-JUL-1999; 99US-0144331P.
PR   27-MAY-1999; 99US-0136392P.     PR   19-JUL-1999; 99US-0144332P.
PR   28-MAY-1999; 99US-0136782P.     PR   19-JUL-1999; 99US-0144333P.
PR   01-JUN-1999; 99US-0137222P.     PR   19-JUL-1999; 99US-0144334P.
PR   03-JUN-1999; 99US-0137528P.     PR   19-JUL-1999; 99US-0144335P.
PR   04-JUN-1999; 99US-0137502P.     PR   20-JUL-1999; 99US-0144352P.
PR   07-JUN-1999; 99US-0137724P.     PR   20-JUL-1999; 99US-0144632P.
PR   08-JUN-1999; 99US-0138094P.     PR   20-JUL-1999; 99US-0144884P.
PR   10-JUN-1999; 99US-0138540P.     PR   21-JUL-1999; 99US-0144814P.
PR   10-JUN-1999; 99US-0138847P.     PR   21-JUL-1999; 99US-0145086P.
PR   14-JUN-1999; 99US-0139119P.     PR   21-JUL-1999; 99US-0145088P.
PR   16-JUN-1999; 99US-0139452P.     PR   22-JUL-1999; 99US-0145085P.
PR   16-JUN-1999; 99US-0139453P.     PR   22-JUL-1999; 99US-0145087P.
PR   17-JUN-1999; 99US-0139492P.     PR   22-JUL-1999; 99US-0145089P.
PR   18-JUN-1999; 99US-0139454P.     PR   22-JUL-1999; 99US-0145192P.
                                     PR   23-JUL-1999; 99US-0145145P.
                                     PR   23-JUL-1999; 99US-0145218P.
                                     PR   23-JUL-1999; 99US-0145224P.
                                     PR   26-JUL-1999; 99US-0145276P.
                                     PR   27-JUL-1999; 99US-0145913P.
                                     PR   27-JUL-1999; 99US-0145918P.
                                     PR   27-JUL-1999; 99US-0145919P.
                                     PR   28-JUL-1999; 99US-0145951P.
                                     PR   02-AUG-1999; 99US-0146386P.
                                     PR   02-AUG-1999; 99US-0146388P.
                                     PR   02-AUG-1999; 99US-0146389P.
                                     PR   03-AUG-1999; 99US-0147038P.
                                     PR   04-AUG-1999; 99US-0147204P.
                                     PR   04-AUG-1999; 99US-0147302P.
                                     PR   05-AUG-1999; 99US-0147192P.
                                     PR   05-AUG-1999; 99US-0147260P.
                                     PR   06-AUG-1999; 99US-0147303P.
                                     PR   06-AUG-1999; 99US-0147416P.
                                     PR   09-AUG-1999; 99US-0147493P.
                                     PR   09-AUG-1999; 99US-0147935P.
                                     PR   10-AUG-1999; 99US-0148171P.
                                     PR   11-AUG-1999; 99US-0148319P.
                                     PR   12-AUG-1999; 99US-0148341P.
                                     PR   13-AUG-1999; 99US-0148565P.
                                     PR   13-AUG-1999; 99US-0148884P.
                                     PR   16-AUG-1999; 99US-0149368P.
```

Db 94 ATGAGCAGTCGTTCGAGTAGAACCGTGTACGTCGGAAACCTTCCTGGCGATATCCGTGAG 153

Qy 21 CysGluValGluAspLeuPheTyrLysGlyProIleValArgTyrGlyIleAspLeuLysIle 40
        |||:::||||||||||
Db 154 AGAGAGGTCGAAGATTTGTTCAGTAAGTATGGACCTGTTGTTCAAATTGATTTGAAGTT 213

Qy 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
        |||||||||||||||:::|||
Db 214 CCTCCAAGGCCTCCTGGTTATGCATTCGTTGAGTTTGAAGATGCTCGGGATCCTGAAGAT 273

Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
        |||||:::||
Db 274 GCTATTCATGGTCGTGATGGCTATGACTTTGATGGGCATCGTTTGAGGGTGGAATTCGCG 333

Qy 81 HisGlyGlyArgPheSerProSerValAspArgTyrSerSerTyrSer----- 98
        |||:::|
Db 334 CATGTTGGGAGGCGT-------TCATCAGATGATACTCGGGGTAGTTTCAATGGTGT 384

Qy 99 -------------------------AlaSerArgAlaProSerArgArgSer 107
Db 385 GGCCGTGGTGGTGGTCGCGGCCGCTGGTGATGGTGGTAGTCGTCGTGGCCATCTAGGAGATCA 444

Qy 108 AspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAsp 127
        |||:::|||
Db 445 GAGTTTCGTGTTCTAGTCACAGGTCTTGCCTTCATCTGCTTCCTTGGCAAGATCTCAAGGAT 504

Qy 128 HisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMet 147
        ||||||:::|||
Db 505 CACATGCGTAAAGGAGGAGGCGATGTCTGTTTCTGCAAGTGTACCGTGATGCTAGGAGGACA 564

Qy 148 SerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAsp 167
        :::|||
Db 565 ACTGGAGTTGTTGATTACACCTGCTATGGACGACCATGAAGTAGTGCGCTGAAAAGCTCGAC 624

Qy 168 AlaThrGluPheArgAsnAlaPheSerGluAlaTyrIleArgValArgGluTyrGluSer 187
        |||||:::|||
Db 625 GACACAGAGTTTCGAAATGCGTTTTCGAATGGAATGTCCGGGTTAGAGAATATGATTCA 684

Qy 188 ArgSerValSerArgSerProAspAspSerLysSerTyr---ArgSerArgSerArgSer 206
        :::|||||||
Db 685 AGGAAGGATTCTAGGAGTCCTAGCCGGGAAGATCCTATTCTAAGAGCCGCAGCCGCAGC 744

Qy 207 ArgGlyProSerCysSerTyrSerSerArgSerValSerProAlaArgSerIle 226
        |||:::|||
Db 745 ---------------AGCCCGAAGCGAGGAGGAGCAGGAGCAGGAGCAGGAGCAGAGA 795

Qy 227 SerProArgSerArgProLeuSerArgArgSerLeuTyrSerSerValSerArgSer 246
        |||:::|||
Db 796 AGTCCCAAGGCAAAG------TCTTCACGTAGGTCCCCGCAAAATCTACATCAAGATCT 849

Qy 247 GlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSer-----Arg 264
        |||:::|
Db 850 CCTGGCCCCGC----------------TCGAAGTCAAGGTCACCGTCTCCAAGAGA 891

Qy 265 SerArgSerArgSerAsnSerProValSerProVal 276
        ||||||:::
Db 892 TCGCGTTCAAGATCAAGATCTCCTCTACCTTCTGTT 927

RESULT 8
AAC38528
ID AAC38528 standard; DNA; 1190 BP.
XX
AC AAC38528;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 21293.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.

PR 99US-0149175P. 17-AUG-1999;
PR 99US-0149426P. 18-AUG-1999;
PR 99US-0149722P. 20-AUG-1999;
PR 99US-0149723P. 20-AUG-1999;
PR 99US-0149929P. 20-AUG-1999;
PR 99US-0149902P. 23-AUG-1999;
PR 99US-0149930P. 23-AUG-1999;
PR 99US-0150566P. 25-AUG-1999;
PR 99US-0150884P. 26-AUG-1999;
PR 99US-0150065P. 27-AUG-1999;
PR 99US-0151066P. 27-AUG-1999;
PR 99US-0151080P. 27-AUG-1999;
PR 99US-0151303P. 30-AUG-1999;
PR 99US-0151438P. 31-AUG-1999;
PR 99US-0151930P. 01-SEP-1999;
PR 99US-0152363P. 07-SEP-1999;
PR 99US-0153070P. 10-SEP-1999;
PR 99US-0153758P. 13-SEP-1999;
PR 99US-0154018P. 15-SEP-1999;
PR 99US-0154039P. 16-SEP-1999;
PR 99US-0154779P. 20-SEP-1999;
PR 99US-0155139P. 22-SEP-1999;
PR 99US-0154486P. 23-SEP-1999;
PR 99US-0155659P. 24-SEP-1999;
PR 99US-0156458P. 28-SEP-1999;
PR 99US-0156596P. 29-SEP-1999;
PR 99US-0157117P. 04-OCT-1999;
PR 99US-0157753P. 05-OCT-1999;
PR 99US-0157865P. 06-OCT-1999;
PR 99US-0158029P. 07-OCT-1999;
PR 99US-0158232P. 08-OCT-1999;
PR 99US-0158369P. 12-OCT-1999;
PR 99US-0152293P. 13-OCT-1999;
PR 99US-0152294P. 13-OCT-1999;
PR 99US-0159295P. 13-OCT-1999;
PR 99US-0159329P. 14-OCT-1999;
PR 99US-0159330P. 14-OCT-1999;
PR 99US-0159331P. 14-OCT-1999;
PR 99US-0159637P. 14-OCT-1999;
PR 99US-0159638P. 14-OCT-1999;
PR 99US-0159584P. 18-OCT-1999;
PR 99US-0160741P. 21-OCT-1999;
PR 99US-0160767P. 21-OCT-1999;
PR 99US-0160768P. 21-OCT-1999;
PR 99US-0160770P. 21-OCT-1999;
PR 99US-0160814P. 21-OCT-1999;
PR 99US-0160815P. 21-OCT-1999;
PR 99US-0160980P. 22-OCT-1999;
PR 99US-0160981P. 22-OCT-1999;
PR 99US-0160989P. 22-OCT-1999;
PR 99US-0161404P. 25-OCT-1999;
PR 99US-0161405P. 25-OCT-1999;
PR 99US-0161406P. 25-OCT-1999;
PR 99US-0161359P. 26-OCT-1999;
PR 99US-0161360P. 26-OCT-1999;
PR 99US-0161361P. 26-OCT-1999;
PR 99US-0161920P. 28-OCT-1999;
PR 99US-0161992P. 28-OCT-1999;
PR 99US-0161993P. 29-OCT-1999;
PR 99US-0162142P. 29-OCT-1999;

Alignment Scores:
Pred. No.:   1.38e-56   Length: 1232
Score: 865.00   Matches: 184
Percent Similarity: 74.32%   Conservative: 33
Best Local Similarity: 63.01%   Mismatches: 45
Query Match: 59.61%   Indels: 30
DB: 3   Gaps: 7

US-10-014-927-19 (1-279) x AAC34814 (1-1232)

Qy 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20

```
XX    EP1033405-A2.
PN
XX
PD    06-SEP-2000.
XX
PF    25-FEB-2000; 2000EP-00301439.
XX

PR    25-FEB-1999;  99US-0121825P.      PR    01-JUL-1999;  99US-0141842P.
PR    05-MAR-1999;  99US-0123180P.      PR    01-JUL-1999;  99US-0142154P.
PR    09-MAR-1999;  99US-0123548P.      PR    02-JUL-1999;  99US-0142055P.
PR    23-MAR-1999;  99US-0125788P.      PR    06-JUL-1999;  99US-0142390P.
PR    25-MAR-1999;  99US-0126264P.      PR    08-JUL-1999;  99US-0142803P.
PR    29-MAR-1999;  99US-0126785P.      PR    09-JUL-1999;  99US-0142920P.
PR    01-APR-1999;  99US-0127462P.      PR    12-JUL-1999;  99US-0142977P.
PR    06-APR-1999;  99US-0128234P.      PR    13-JUL-1999;  99US-0143542P.
PR    08-APR-1999;  99US-0128714P.      PR    14-JUL-1999;  99US-0143624P.
PR    16-APR-1999;  99US-0129845P.      PR    15-JUL-1999;  99US-0144005P.
PR    19-APR-1999;  99US-0130077P.      PR    16-JUL-1999;  99US-0144085P.
PR    21-APR-1999;  99US-0130449P.      PR    16-JUL-1999;  99US-0144086P.
PR    23-APR-1999;  99US-0130510P.      PR    19-JUL-1999;  99US-0144325P.
PR    28-APR-1999;  99US-0130891P.      PR    19-JUL-1999;  99US-0144331P.
PR    28-APR-1999;  99US-0111449P.      PR    19-JUL-1999;  99US-0144332P.
PR    30-APR-1999;  99US-0132048P.      PR    19-JUL-1999;  99US-0144333P.
PR    30-APR-1999;  99US-0132407P.      PR    19-JUL-1999;  99US-0144334P.
PR    04-MAY-1999;  99US-0132484P.      PR    19-JUL-1999;  99US-0144335P.
PR    05-MAY-1999;  99US-0132485P.      PR    20-JUL-1999;  99US-0144352P.
PR    06-MAY-1999;  99US-0132486P.      PR    20-JUL-1999;  99US-0144632P.
PR    07-MAY-1999;  99US-0132863P.      PR    20-JUL-1999;  99US-0144884P.
PR    11-MAY-1999;  99US-0134256P.      PR    21-JUL-1999;  99US-0144814P.
PR    14-MAY-1999;  99US-0134218P.      PR    21-JUL-1999;  99US-0145086P.
PR    14-MAY-1999;  99US-0134219P.      PR    21-JUL-1999;  99US-0145088P.
PR    14-MAY-1999;  99US-0134221P.      PR    22-JUL-1999;  99US-0145085P.
PR    14-MAY-1999;  99US-0134370P.      PR    22-JUL-1999;  99US-0145087P.
PR    18-MAY-1999;  99US-0134768P.      PR    22-JUL-1999;  99US-0145089P.
PR    19-MAY-1999;  99US-0134941P.      PR    22-JUL-1999;  99US-0145192P.
PR    20-MAY-1999;  99US-0135124P.      PR    23-JUL-1999;  99US-0145145P.
PR    21-MAY-1999;  99US-0135353P.      PR    23-JUL-1999;  99US-0145218P.
PR    24-MAY-1999;  99US-0135629P.      PR    23-JUL-1999;  99US-0145224P.
PR    25-MAY-1999;  99US-0136021P.      PR    26-JUL-1999;  99US-0145276P.
PR    27-MAY-1999;  99US-0136392P.      PR    27-JUL-1999;  99US-0145913P.
PR    28-MAY-1999;  99US-0136782P.      PR    27-JUL-1999;  99US-0145918P.
PR    01-JUN-1999;  99US-0137222P.      PR    27-JUL-1999;  99US-0145919P.
PR    03-JUN-1999;  99US-0137528P.      PR    28-JUL-1999;  99US-0145951P.
PR    04-JUN-1999;  99US-0137502P.      PR    02-AUG-1999;  99US-0146386P.
PR    07-JUN-1999;  99US-0137724P.      PR    02-AUG-1999;  99US-0146388P.
PR    08-JUN-1999;  99US-0138094P.      PR    02-AUG-1999;  99US-0146389P.
PR    10-JUN-1999;  99US-0138540P.      PR    03-AUG-1999;  99US-0147038P.
PR    10-JUN-1999;  99US-0138847P.      PR    04-AUG-1999;  99US-0147204P.
PR    14-JUN-1999;  99US-0139119P.      PR    04-AUG-1999;  99US-0147302P.
PR    16-JUN-1999;  99US-0139452P.      PR    05-AUG-1999;  99US-0147192P.
PR    16-JUN-1999;  99US-0139453P.      PR    05-AUG-1999;  99US-0147260P.
PR    17-JUN-1999;  99US-0139492P.      PR    06-AUG-1999;  99US-0147303P.
PR    18-JUN-1999;  99US-0139454P.      PR    06-AUG-1999;  99US-0147416P.
PR    18-JUN-1999;  99US-0139455P.      PR    09-AUG-1999;  99US-0147493P.
PR    18-JUN-1999;  99US-0139456P.      PR    09-AUG-1999;  99US-0147935P.
PR    18-JUN-1999;  99US-0139457P.      PR    10-AUG-1999;  99US-0148171P.
PR    18-JUN-1999;  99US-0139458P.      PR    11-AUG-1999;  99US-0148319P.
PR    18-JUN-1999;  99US-0139459P.      PR    12-AUG-1999;  99US-0148341P.
PR    18-JUN-1999;  99US-0139460P.      PR    13-AUG-1999;  99US-0148565P.
PR    18-JUN-1999;  99US-0139461P.      PR    13-AUG-1999;  99US-0148684P.
PR    18-JUN-1999;  99US-0139462P.      PR    16-AUG-1999;  99US-0149368P.
PR    18-JUN-1999;  99US-0139463P.      PR    17-AUG-1999;  99US-0149175P.
PR    18-JUN-1999;  99US-0139750P.      PR    18-AUG-1999;  99US-0149426P.
PR    18-JUN-1999;  99US-0139763P.      PR    20-AUG-1999;  99US-0149722P.
PR    21-JUN-1999;  99US-0139817P.      PR    20-AUG-1999;  99US-0149723P.
PR    21-JUN-1999;  99US-0139899P.      PR    20-AUG-1999;  99US-0149929P.
PR    23-JUN-1999;  99US-0140353P.      PR    23-AUG-1999;  99US-0149902P.
PR    23-JUN-1999;  99US-0140354P.      PR    25-AUG-1999;  99US-0149930P.
PR    24-JUN-1999;  99US-0140695P.      PR    26-AUG-1999;  99US-0150556P.
PR    28-JUN-1999;  99US-0140823P.      PR    27-AUG-1999;  99US-0150884P.
PR    29-JUN-1999;  99US-0140991P.      PR    27-AUG-1999;  99US-0151065P.
PR    30-JUN-1999;  99US-0141287P.      PR    30-AUG-1999;  99US-0151066P.
                                         PR    31-AUG-1999;  99US-0151080P.
                                         PR    01-SEP-1999;  99US-0151303P.
                                         PR    07-SEP-1999;  99US-0151438P.
                                         PR    10-SEP-1999;  99US-0151930P.
                                         PR    13-SEP-1999;  99US-0152363P.
                                         PR    15-SEP-1999;  99US-0153070P.
                                         PR                  99US-0153758P.
                                         PR                  99US-0154018P.
```

PR   16-SEP-1999;   99US-0154039P.
PR   20-SEP-1999;   99US-0154779P.
PR   22-SEP-1999;   99US-0155139P.
PR   23-SEP-1999;   99US-0155486P.
PR   24-SEP-1999;   99US-0155659P.
PR   28-SEP-1999;   99US-0156458P.
PR   29-SEP-1999;   99US-0156596P.
PR   04-OCT-1999;   99US-0157117P.
PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157865P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158369P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   29-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

Alignment Scores:
Pred. No.:
Score: 863.00
Percent Similarity: 73.31%
Best Local Similarity: 62.16%
Query Match: 59.48%
DB: 3

Length: 1190
Matches: 184
Conservative: 33
Mismatches: 45
Indels: 34
Gaps: 7

1.88e-56

US-10-014-927-19 (1-279) x AAC38528 (1-1190)

Qy   1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db   59 ATGAGCAGTCGTTCGAGTAGAACCGTGTACGTCGGAAACCTTCCTGGCGATATCCGTGAG 118

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyTyrGlyProIleValAspIleLeuLysIle 40
Db 119 AGAGAGGTCGAAGATTGTTCAGTAAGTAGGACCTGTGTTGTTCAAATTGATTTGAAGGTT 178

Qy   41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 179 CCTCCAAGGCCTCCTGGTTATGCATTCGTTGAGTTTGATGATCCTCGGGATGCTGAAGAT 238

Qy   61 AlaIleTyrArgAspGlyTyrTyrArgSerPheAspGlyCysArgLeuArgValGluIleAla 80
Db 239 GCTATTCATGGTCGTGATGGCTATGACTTTGATGGGCATCGTTTGAGGGTGGAATTGGCG 298

Qy   81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSer------ 98
Db 299 CATGGTGGAGGGCGT--------TCATCAGATGATACTCGGGTAGTTTCAATGGTGGT 349

Qy   99 -------------AlaSerArgAlaPro 103

Db 350 GGCCGTGGTGGTGGTCGGGACCCGTCGGTGGATGGCGCCGTCGTCGTCGGTGGATGGTGGTGGTAGTCGTGGGGCCA 409

Qy 104 SerArgArgSerAspTyrArgGlyValLeuValThrGlyValLeuProProSerAlaSerTrpGln 123
Db 410 TCTAGGAGGATCAGAGTTTCGTGTTCTAGTCACAGGCTGCCTTCATCGCTTCTTGGCAA 469

Qy 124 AspLeuLysAspHisMetArgLysAlaGlyAlaLysAspValCysPheSerGluValPheProAsp 143
Db 470 GATCTCAAGGATCACATGCGTAAAGAGGAGGCGATGTCGTTTCTCGCAAGTGTACCGTGAT 529

Qy 144 ArgLysGlyMetSerGlyValValValAspTyrSerAsnTyrAspMetLysTyrAlaIle 163
Db 530 GCTAGAGGGACAACTGGAGTTGTTGATTACACCTGCTATGAGGACATGAAGTATGCGCTG 589

Qy 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerArgAlaTyrIleArgValArg 183
Db 590 AAAAAGCTCGACGACACAGAGTTTCGAAATGCGTTTTCGAATGGATATGTCCGGGTTAGA 649

Qy 184 GluTyrGluSerArgSerArgSerProAspAspSerLysSerTyr---ArgSer 202
Db 650 GAATATGATTCAAGGAAGGATTCTAGGAGTCCTAGCCGGGGAGATCCTATTCTAAGAGC 709

Qy 203 ArgSerArgSerArgGlyProSerSerTyrSerSerLysSerArgSerValSerPro 222
Db 710 CGGAGCCCGCAGCCTGGACGAAGCCTG------AGCCGGAGCAGGAGCAGAAGCAGG 760

Qy 223 AlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSer 242
Db 761 AGCAGGAGCAGAAGTCCCAAGGCAAAG------TCTTCACGTAGTCCCCTGCAAAATCT 814

Qy 243 ValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLys 262
Db 815 ACATCAAGGATCTCCTGGCCCCGC--------------TCGAAGTCAAGGTCACCG 856

Qy 263 Ser------ArgSerArgSerSerAsnSerProValSerPro 276
Db 857 TCTCCAAGAAGATCGCGTTCAAGATCAAGAATCCTCACCTTCTGTT 904

RESULT 9
AAC42654
ID   AAC42654 standard; DNA; 885 BP.
XX
AC   AAC42654;
XX
DT   17-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 36361.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000BP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.

PR 23-APR-1999; 99US-0105510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0137782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0180094P.
PR 10-JUN-1999; 99US-0185040P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0191119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0194492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0194462P.
PR 18-JUN-1999; 99US-0194463P.
PR 18-JUN-1999; 99US-0197750P.
PR 18-JUN-1999; 99US-0197763P.
PR 21-JUN-1999; 99US-0198817P.
PR 22-JUN-1999; 99US-0198899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140695P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142154P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 13-JUL-1999; 99US-0142977P.
PR 14-JUL-1999; 99US-0143542P.
PR 15-JUL-1999; 99US-0143624P.
PR 16-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

```
 184 GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr---ArgSer 202
     |||:||||  :::||||                    ::|||||||            :::||
 591 GAATATGATTCAAGGAGGAGGATTCGAGGAGATCCAGCCGTGGAAGATCCTATTCTAAGAGC 650

 203 ArgSerArgSerArgTyrProSerCysSerSerTyrSerSerLysSerProSerValSerPro 222
     |||:|||                               ::||||||||
 651 CGCAGCCGTGGACGGAGTCCAGC----------CGTAGTCGTAGCCGCAGCAGAAGCAGG 701

 223 AlaArgSerIleSerProArgProLeuSerArgSerArgLeuTyrSerSer 242
     :::::|||  |||::||:::|||  ::||||::  :::
 702 AGCAAGAGCAGAGTCCAAAGGCTAAATCTTTGCGTAGATCGCCTGCA------AAATCT 755

 243 ValSerArgSerGlySerLeuLeuArgAlaAlaGlyAspTrpIleSerGlnSerArgSerLys 262
     ||||||                                       ||:::
 756 ACATCGGAGATCT--------------------------------CCTCGCTCTCGC 779

 263 SerArgSerArgSerArgSerAsnSerPro 272
     ||||:::::::|||
 780 TCCGCTCAAGTCGAGGTCACTGTCTCCA 809
```

RESULT 10
AAC37026
ID   AAC37026   standard; DNA; 1241 BP.
XX
AC   AAC37026;
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
DT   17-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 15916.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX

Alignment Scores:
Pred. No.:   9.62e-55
Score: 838.50
Percent Similarity: 72.07%
Best Local Similarity: 62.07%
Query Match: 57.79%
DB:

```
Length:        885
Matches:       180
Conservative:  29
Mismatches:    43
Indels:        39
Gaps:           7
```

US-10-014-927-19 (1-279) x AAC42654 (1-885)

```
Qy    1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys  20
          :::||||  |||::|||
Db    1 ATGAGCAGCCGTTCGAGTAGAACGATTTACGTCGGGAACCTTCCCGGCGATATCCGTGAA  60

Qy   21 CysGlyValGluAspPheTyrLysTyrGlyTyrGlyProIleValAspIleAspLeuLysIle  40
           ||||                               ::|||
Db   61 AGAGAAGTTGAAGACTTGTTCAGTAAGTAAGTATGGACCTGTTGTTCAAATCGATTTGAAGATT 120

Qy   41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp  60
          |||||||  |||  |||  ||||||  :::|||||||  |||::||||
Db  121 CCGCCGAGGCCCTCCCAGGCTATGGCATTCGTCGTAGTCGTGATGCTCGTGATGCTGATGAT 180

Qy   61 AlaIleTyrGlyIleArgAspGlyTyrArgAspPheAspGlyTyrCysArgLeuArgValGluIleAla  80
          ||||||                               ::|||||
Db  181 GCAATTTATGGCCGTGATGGTTATGACTTTGATGGGCATCATTTACGGTGGAACTAGCT 240

Qy   81 HisGlyLysArgArgPheSerProSerValAspArgTyrSerSerTyrSerAla---  99
          ||||                               ||||||
Db  241 CATGGTGGGAGGCGTTCATCACATGATGCA--------CGCGGTAGTTATAGTGGTCGT 291

Qy  100 ----------SerArgAlaProSerArgArg 106
                    |||  ::|||||||
Db  292 GGTCGTGCGCGGTCGTCGTGGTGGTGGTGACGGTGGTGCTGAACGTGGACCATCTAGGAGA 351

Qy  107 SerAspTyrArgValLeuValThrGlyValGlyAspValCysPheSerAlaSerTrpGlnAspLeuLys 126
          ||||||                               ::||||||
Db  352 TCAGAGTACCCGCGTTGTGTGTAGTGTCAGGTTGCCTTCATCTGGCCTCCTGGCAAGACCTCAAG 411

Qy  127 AspHisMetArgLysAlaGlyAlaGlyValCysPheSerGluValPheProAspArgLysGly 146
          |||  ::|||  |||  ::|||||||  ||||||::
Db  412 GATCACATGCGTAAAGCAGGAGCAGAAGTTTGTTTTTCTGATGTGTTCCGTGATGGTAGAGGT 471

Qy  147 MetSerGlyValValAspTyrSerArgAsnTyrAspMetLysTyrAlaIleArg------ 164
          ||||||                               :::
Db  472 ACAACTGGAATTGTAGATTATACCAGCTACGAGGACACTAGACGGGTGAG-ATATAA 530

Qy  165 ---LysLeuAspAlaThrAspThrCysCysArgAsnAlaPheSerSerArgAlaTyrIleArgValArg 183
          :::||::::::||||                    ||||||
Db  531 AAAAAGCTCGATGACACAGAGTTTCGGAATGCGTTTTCCATGAATATGTTCGGGGTTAGA 590
```

```
PR  24-MAY-1999;  99US-0135629P.
PR  25-MAY-1999;  99US-0136021P.
PR  27-MAY-1999;  99US-0136392P.
PR  28-MAY-1999;  99US-0167782P.
PR  01-JUN-1999;  99US-0137222P.
PR  03-JUN-1999;  99US-0137528P.
PR  04-JUN-1999;  99US-0137502P.
PR  07-JUN-1999;  99US-0137724P.
PR  08-JUN-1999;  99US-0138094P.
PR  10-JUN-1999;  99US-0138540P.
PR  10-JUN-1999;  99US-0138847P.
PR  14-JUN-1999;  99US-0139119P.
PR  16-JUN-1999;  99US-0139452P.
PR  16-JUN-1999;  99US-0139453P.
PR  17-JUN-1999;  99US-0139492P.
PR  18-JUN-1999;  99US-0139454P.
PR  18-JUN-1999;  99US-0139455P.
PR  18-JUN-1999;  99US-0139456P.
PR  18-JUN-1999;  99US-0139457P.
PR  18-JUN-1999;  99US-0139458P.
PR  18-JUN-1999;  99US-0139459P.
PR  18-JUN-1999;  99US-0139460P.
PR  18-JUN-1999;  99US-0139461P.
PR  18-JUN-1999;  99US-0139462P.
PR  18-JUN-1999;  99US-0139463P.
PR  18-JUN-1999;  99US-0139750P.
PR  18-JUN-1999;  99US-0139763P.
PR  21-JUN-1999;  99US-0139817P.
PR  22-JUN-1999;  99US-0139899P.
PR  23-JUN-1999;  99US-0140353P.
PR  01-JUN-1999;  99US-0140354P.
PR  24-JUN-1999;  99US-0140695P.
PR  28-JUN-1999;  99US-0140823P.
PR  29-JUN-1999;  99US-0140991P.
PR  30-JUN-1999;  99US-0141287P.
PR  01-JUL-1999;  99US-0141842P.
PR  01-JUL-1999;  99US-0142154P.
PR  02-JUL-1999;  99US-0142055P.
PR  06-JUL-1999;  99US-0142390P.
PR  08-JUL-1999;  99US-0142803P.
PR  09-JUL-1999;  99US-0142920P.
PR  12-JUL-1999;  99US-0142977P.
PR  13-JUL-1999;  99US-0143542P.
PR  14-JUL-1999;  99US-0143624P.
PR  15-JUL-1999;  99US-0144005P.
PR  16-JUL-1999;  99US-0144085P.
PR  16-JUL-1999;  99US-0144086P.
PR  19-JUL-1999;  99US-0144325P.
PR  19-JUL-1999;  99US-0144331P.
PR  19-JUL-1999;  99US-0144332P.
PR  19-JUL-1999;  99US-0144333P.
PR  19-JUL-1999;  99US-0144334P.
PR  19-JUL-1999;  99US-0144335P.
PR  20-JUL-1999;  99US-0144352P.
PR  20-JUL-1999;  99US-0144332P.
PR  21-JUL-1999;  99US-0144884P.
PR  21-JUL-1999;  99US-0144814P.
PR  21-JUL-1999;  99US-0145086P.
PR  22-JUL-1999;  99US-0145088P.
PR  22-JUL-1999;  99US-0145085P.
PR  22-JUL-1999;  99US-0145087P.
PR  22-JUL-1999;  99US-0145089P.
PR  22-JUL-1999;  99US-0145192P.
PR  23-JUL-1999;  99US-0145145P.
PR  23-JUL-1999;  99US-0145218P.
PR  23-JUL-1999;  99US-0145224P.
PR  26-JUL-1999;  99US-0145276P.
PR  27-JUL-1999;  99US-0145913P.
PR  27-JUL-1999;  99US-0145918P.
PR  27-JUL-1999;  99US-0145919P.
PR  28-JUL-1999;  99US-0145951P.
PR  02-AUG-1999;  99US-0146386P.
PR  02-AUG-1999;  99US-0146388P.

PR  02-AUG-1999;  99US-0146389P.
PR  03-AUG-1999;  99US-0147038P.
PR  04-AUG-1999;  99US-0147204P.
PR  04-AUG-1999;  99US-0147302P.
PR  05-AUG-1999;  99US-0147192P.
PR  05-AUG-1999;  99US-0147260P.
PR  06-AUG-1999;  99US-0147303P.
PR  06-AUG-1999;  99US-0147416P.
PR  09-AUG-1999;  99US-0147493P.
PR  09-AUG-1999;  99US-0147935P.
PR  10-AUG-1999;  99US-0148171P.
PR  11-AUG-1999;  99US-0148319P.
PR  12-AUG-1999;  99US-0148341P.
PR  13-AUG-1999;  99US-0148565P.
PR  13-AUG-1999;  99US-0148684P.
PR  16-AUG-1999;  99US-0149368P.
PR  17-AUG-1999;  99US-0149175P.
PR  18-AUG-1999;  99US-0149426P.
PR  20-AUG-1999;  99US-0149722P.
PR  20-AUG-1999;  99US-0149723P.
PR  20-AUG-1999;  99US-0149929P.
PR  23-AUG-1999;  99US-0149902P.
PR  23-AUG-1999;  99US-0149930P.
PR  25-AUG-1999;  99US-0150566P.
PR  26-AUG-1999;  99US-0150884P.
PR  27-AUG-1999;  99US-0151065P.
PR  27-AUG-1999;  99US-0151066P.
PR  27-AUG-1999;  99US-0151080P.
PR  30-AUG-1999;  99US-0151303P.
PR  31-AUG-1999;  99US-0151438P.
PR  01-SEP-1999;  99US-0151930P.
PR  07-SEP-1999;  99US-0152263P.
PR  10-SEP-1999;  99US-0153070P.
PR  13-SEP-1999;  99US-0153758P.
PR  15-SEP-1999;  99US-0154018P.
PR  16-SEP-1999;  99US-0154039P.
PR  20-SEP-1999;  99US-0154779P.
PR  22-SEP-1999;  99US-0155139P.
PR  23-SEP-1999;  99US-0155486P.
PR  24-SEP-1999;  99US-0155659P.
PR  28-SEP-1999;  99US-0156458P.
PR  29-SEP-1999;  99US-0156596P.
PR  04-OCT-1999;  99US-0157117P.
PR  05-OCT-1999;  99US-0157753P.
PR  06-OCT-1999;  99US-0157865P.
PR  07-OCT-1999;  99US-0158029P.
PR  08-OCT-1999;  99US-0158232P.
PR  12-OCT-1999;  99US-0158369P.
PR  13-OCT-1999;  99US-0159293P.
PR  13-OCT-1999;  99US-0159294P.
PR  14-OCT-1999;  99US-0159295P.
PR  14-OCT-1999;  99US-0159329P.
PR  14-OCT-1999;  99US-0159330P.
PR  14-OCT-1999;  99US-0159331P.
PR  14-OCT-1999;  99US-0159637P.
PR  18-OCT-1999;  99US-0159638P.
PR  21-OCT-1999;  99US-0159584P.
PR  21-OCT-1999;  99US-0160741P.
PR  21-OCT-1999;  99US-0160767P.
PR  22-OCT-1999;  99US-0160768P.
PR  22-OCT-1999;  99US-0160770P.
PR  25-OCT-1999;  99US-0160814P.
PR  25-OCT-1999;  99US-0160815P.
PR  26-OCT-1999;  99US-0160980P.
PR  26-OCT-1999;  99US-0160981P.
PR  28-OCT-1999;  99US-0160989P.
                  99US-0161404P.
                  99US-0161405P.
                  99US-0161406P.
                  99US-0161359P.
                  99US-0161360P.
                  99US-0161361P.
                  99US-0161920P.
```

```
PR    28-OCT-1999;   99US-0161992P.
PR    28-OCT-1999;   99US-0161993P.
PR    29-OCT-1999;   99US-0162142P.

Alignment Scores:
Pred. No.:                        1241
Score:                    805.00
Percent Similarity:       71.67%    Matches:          183
Best Local Similarity:    62.46%    Conservative:      27
Query Match:              55.48%    Mismatches:        53
DB:                       3         Indels:            30
                                    Gaps:              10

US-10-014-927-19 (1-279) x AAC37026 (1-1241)

Qy    1   MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   20
              |||  ||::| |  ::::|||| | ||| ||  :| :|||::|
Db   141   ATGAGTGGGCGGATTTTCTCGGTCAATCTATGTTGGTAACTTGCCCGGTGACATTAGGAA   200

Qy   21   CysGluValGluSerProLeuPheTyrLysTyrGlyTyrGlyProIleValAspIleAspLeuLysIle   40
              ::::| ||||| |::|:  |||:| |: |:| :|:|||:|
Db  201   CATGAGATTGAAGATATCTTTTACAAGTATGGCCCCATTGTCGATATTGAATTGAAGGTT   260

Qy   41   ProProArgProProGlyTyrTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp   60
              ||| ||:|:| | |  | |  ::| |||:::|||
Db  261   CCACCTCGGCCTCCATGTTATTGCTTTGTTGAGTTTGAGCAATTCTCGGGATGCTGAAGAT   320

Qy   61   AlaIleTyrGlyArgAspGlyTyrArgAspPheAspGlyCysAspGlyLeuArgValGluIleAla   80
              ||||| ||||||:::|| | :| |||::|:|| |::|||
Db  321   GCCATCAAGGGCCGTGATGGGCTATAATTGGATGGCTGTCCTTGAGGTTGAGCTTGCA   380

Qy   81   HisGlyGlyArgArgPheSerProSerValAspArgTyr------------Ser   94
              ||:||| | ||| |::|      ||||  | :    ----------
Db  381   CATGGTTGGTCCAGGAGACAGTCTTCAAGT---GATCGTCGTGGTGGTCGTGGTGGTGGC   437

Qy   95   SerSerTyr--------------------------SerAlaSerArgAlaProSerArgArg   106
              ||||:                   | :|:| ::||:
Db  438   AGCGGCTATGGTGGTCGGTGGTGGTCGTGGTCAGCTCGGTTGGCGTCTCACGACAC   497

Qy  107   SerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys   126
              ||:::| :|: |:||||||:|::| |||  ||| ::|||:|
Db  498   TCTGAATTCCGAGTTATTGTACGTGGGCTCCCATCATCTGCTTCATGGCAAGATTTGAAG   557

Qy  127   AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly   146
              |::||: ||| | | ||| |::|  | :|:| | |:: |||:|
Db  558   GATCATATGCGGAAAGCTGGTGATGTGTGCTTTGCTGAGGTGACTCGAGACAGTGATGGA   617

Qy  147   MetSerGlyValValAspTyrSerAsnTyrArgAspMetLysTyrAlaIleArgLeu   166
              ||||||  | | :| |||::|  |::||||| |:|| ||||:|
Db  618   ACTTATGCGTGTTGTCGACTACACCAATTATGATGAAGCATGCAATAAGGAAACTT   677

Qy  167   AspAlaThrGluPheArgGluAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu   186
              |::|::   :  ::| |:::|: |||::|:|||::|||
Db  678   GATGACACAGAGTTCAGAAACCCCTGGGCTAGAGGTTTATCCGGGTTAAGAAATATGAA   737

Qy  187   --SerArgSerValSerArgProAspAspSerLysSerTyrArgSerArgSerArg   205
                --|||||: |: || |||||::|::|:| |:|:|||:|| |
Db  738   AGCTCCCGATCAAGAAGCAGAAGAGCCCAAGCAGCAGAAGT---CGAAGTCGTAGCCGA   794

Qy  206   SerArgGlyProSerCysSerTyrSer--SerLysSerArgSerValSerProAlaArg   224
              ||||||  ||    |: :: :  --|||||| | ||| ::|
Db  795   GGCCGTGGTGGTCGCAGCCATAGCCGCCGGATAGCCCAGCCTTAGCAGGAGCAAGAGGGCCCAAGGAAG   854

Qy  225   SerIle------SerProArg--------SerArgProLeuSerArgSerArgSerLeu   239
              |:|:   ------||||||   --------:||| |: |||
Db  855   GATCTGAGTAAATCACCCAGGCGATCCCTTTCCAGATCGATTTCAAAATCTAGATCG---   911

Qy  240   TyrSerSerValSerArgGlySerLeuLeuArgAlaGlyIleAspTrpIleSerGlnSer   259
              |:: ||| ||||||  |:||:|::| | |||::||
Db  912   ---CCATCCCGACAACAGAAGAAGAGTCCCCCCAGGGCA--------------ATGTCG   953

Qy  260   ArgSerLysSerArgSerArgSerArgSerAsnSerPro   272
              ||||||||| |||| ||| |:|::|::||::|||
Db  954   AGATCAAGTCCCAGGTCCAGGTCCAGATCGAGTCTCG   992
```

```
Alignment Scores:
Pred. No.:                        2765
Score:                    7.13e-35   Matches:           134
Percent Similarity:       584.50     Conservative:       29
Best Local Similarity:    69.07%     Mismatches:         60
Query Match:              56.78%     Indels:             13
DB:                       40.28%     Gaps:                7
                          9

US-10-014-927-19 (1-279) x ADD22452 (1-2765)

Qy    9   IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr   28
              ::   |:|| | | | ||        ::::||:|:::
Db  142   ATCTACGTGGGTAACTTACCTCCAGACATCCGACCAAGACATTGAGGACGTGTTCTAC   201
```

```
XX
DR   WPI; 2001-266308/27.
DR   P-PSDB; AAB90781.
XX
PT   DNA sequences, proteins encoded by them and antibodies against them
PT   useful in diagnosis and treatment of vascular disease caused by
PT   arteriosclerosis.
XX
PS   Claim 20; Page 386-388; 678pp; Japanese.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human shear stress response proteins. These are useful in the
CC   diagnosis, treatment and screening of vascular diseases caused by
CC   arteriosclerosis, including heart failure, post-PTCA restenosis and
CC   hypertension
XX
SQ   Sequence 1428 BP; 341 A; 296 C; 400 G; 391 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                        Length:          1428
Score:               4.63e-35     Matches:          134
Percent Similarity:  582.50       Conservative:      28
Best Local Similarity: 64.80%     Mismatches:        47
Query Match:         53.60%       Indels:            41
DB:                  40.14%       Gaps:               8
                     4

US-10-014-927-19 (1-279)  x AAH02904 (1-1428)
```

```
Qy      9  IleTyrValGlyAsnLeuProGlyAspIleArgIleArgLysCysGluValGluAspLeuPheTyr        28
              :::||||  |||||::||  |  |||||||  |::||||:
Db    176  ATCTACGTGGGTAACTTACCTCCAGACATCCGAACCAAGGACATTGGACGTGTTCTAC               235

Qy     29  LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGlyTyrAla                48
              |||||||||| ||||  |||||||||::||  |||::|||||| ||||||
Db    236  AAATACGGCGCTATCCGCGACACTCGACCTCAGAATCGCCGGGGGACCGCCCTTCCC               295

Qy     49  PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrArgGlyArgProGlyTyrAla        68
              ||:|||||||:::|||  ||||||| |||||  ||||||||:::|||||||
Db    296  TTCGTTGAGTTCGAGGAGTTCCTGAAGCCGGTGTATGGTCGGCACGGCTAT                     355

Qy     69  AspPheAspGlyCysPheAspGlyAspAlaAspProArgArgProAlaArg                      84
              |:::||||    |||||  ||||||
Db    356  GATTACGATGGGTACCGTCGGGGTGGAGTTTCCTGAAGGCCGTGGAACAGGCCGA                 415

Qy     85  ------------------------------------------ArgPheSerProSerValAsp         91
                                                         :::||||:
Db    416  GGCGGCGCGCGGGGTGAGGTGCGGAGCTCCCGAGCTCGTATGCGCCC-------------             466

Qy     92  ArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArgVal             111
              |||||||  |||  |
Db    467  ---------------------------CCATCCAGGCGCGTCGAAAACAGAGTG                  493

Qy    112  LeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLys             131
              :::||||||::||||||||:::||||:||
Db    494  GTTGTCTCTGGACTGCCTCCCAGTGCAGGTGGGAAGGATTAAAGGATCACATGCGTGAA             553

Qy    132  AlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyValVal             151
                                                            :::||||||||||
Db    554  GCAGGTGATGTATGTTTATGCTGATGTTTACCGAGATGGC-----------ACTGTGTCGTG        604

Qy    152  AspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPhe             171
              :::||||  |||||::
Db    605  GAGTTGTACGGAAAGAGATATGACCTATGCAGTTCGAAAACTGGATAACACTAAGTTT              664

Qy    172  Arg---AsnAlaPheSerSerAlaTyrIleArgValArgGluSerArgSerVal                  190
              :::
Db    665  AGATCTCATGAGGGAGAAAACTGCCTACATCCGGGTTAAAGTTGATGGGCCCAGAAGTCCA           724

Qy    191  SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgGlyProSer                  210
              |||:::||
Db    725  AGTTATGGAAGATCTCGATCTCGAAGC---CGTAGTCGTAGCAGAGCCGTAGCAGAAGC             781

Qy    211  CysSerTyrSerArgSerLysSerArgSerValSerProAlaArgSerIle--SerProArg          229
```

Db   782 ------------AACAGCAGGAGTCGCGACTCAGTTACTCCCCAAGGAGGAGAGCAGGAGGATCACCACGC 832
           :::||||::::||||||:|||||| |||||| ||||||

Qy   230 SerArgProLeu---SerArgSerArgSer 238
           ||| |||||||

Db   833 TATTCTCCCGTCATAGCAGATCTCGCTCT 862

RESULT 13
AAL60055
ID   AAL60055 standard; cDNA; 1428 BP.
XX
AC   AAL60055;
XX
DT   27-AUG-2003 (first entry)
XX
DE   Human Pc036-2 cDNA.
XX
KW   Human; differentially regulated protein; prevention; therapy; vaccine;
KW   prostate cancer; gene therapy; pre-mRNA splicing factor; Pc036-2;
KW   chromosome 17q21.3-q22; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   125..871
FT   /*tag= a
FT   /product= "Human Pc036-2 protein"
XX
PN   WO2003040331-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US035563.
XX
PR   07-NOV-2001; 2001US-0331041P.
PR   07-NOV-2001; 2001US-0331042P.
PR   18-DEC-2001; 2001US-0342251P.
PR   07-JAN-2002; 2002US-0344791P.
XX
PA   (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI   Sun Z, Li X, Jay G, Kovacs KF, Fan W;
XX
DR   WPI; 2003-449451/42.
DR   P-PSDB; AAO29561.
XX
PT   New polynucleotide for diagnosing, staging, monitoring, prognosticating,
PT   preventing or treating, or determining the predisposition to, diseases or
PT   conditions such as prostate cancer, and for research or forensic science.
XX
PS   Claim 29; Page 127-128; 100pp; English.
XX
CC   The present invention relates to novel differentially regulated genes and
CC   polypeptides encoded by them. Sequences of the invention are useful in
CC   diagnosing, staging, monitoring, prognosticating, preventing, treating or
CC   determining the predisposition to diseases or conditions such as prostate
CC   cancer. They may be used as molecular markers, drug targets, vaccines, in
CC   gene therapy, research, clinical medicine or forensic science. The
CC   present sequence is a differentially regulated prostate cDNA, Pc036-2
CC   which codes for a pre-mRNA splicing factor. Pc036-2 gene is located on
CC   chromosome 17q21.3-q22
XX
SQ   Sequence 1428 BP; 341 A; 296 C; 400 G; 391 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score: 582.50
Percent Similarity: 64.80%
Best Local Similarity: 53.60%
Query Match: 40.14%
DB:

Length: 1428
Matches: 134
Conservative: 28
Mismatches: 47
Indels: 41
Gaps: 8

4.63e-35

US-10-014-927-19 (1-279) x AAL60055 (1-1428)

Qy   9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluValAspLeuPheTyr 28
           ||||||||||||||||:||||| |||||||||||||||||||||||||||:|||||
Db   176 ATCTACGTGGGTAACTTACCTCCAGACATCCGAACCAAGGACATTGAGGACGACGTTCTAC 235

Qy   29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
           :::|||||||||||||||||||||:::||||||||||||||||||||||||||:::
Db   236 AAATACGGACCTATCCCGACATCGACCTCAAGAATCCGCCGCGGGGGACCGCCCTTCGCC 295

Qy   49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
           ||||||||::::|||||||||:||||||:::||||||||:|||||||||:||
Db   296 TTCGTTGAGTTCGAGGACCCGCGAGACGCCGGAAGACACGGGTGTATCGTCGCGACGGCTAT 355

Qy   69 AspPheAspGlyCysArgGlyValGluIleAlaHisGlyArg------------ 84
           ||||||:::|||||||||||||||||||||||||
Db   356 GATTACGATGGGTACCGTCTCGCCGGTGAGTTTCCTCAAGCCGCCGGGGAACAGGCCGA 415

Qy   85 ------------------------ArgPheSerProSerValAsp 91
                                    ||||||:::
Db   416 GGGGGCGCGGGGGTGGAGGTGGCGGAGCTCCCGAGGTCGCTATGGCCCC----------- 466

Qy   92 ArgTyrSerSerTyrSerAlaSerArgArgSerArgSerAspTyrArgVal 111
           ||||||:::||::::
Db   467 -----------------------CCATCCAGGCGGTCTGAAAACAGAGTG 493

Qy   112 LeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLys 131
           :::||:::|||:::|||:||||||||||||||||||||||:::
Db   494 GTTGTTCTCGGACGTGCCTCCAAGTGGAGTTGCAGGATTTAAAGGATCACATGCGTGAA 553

Qy   132 AlaGlyAspValCysPheSerGluValPheProSerGluLysGlyMetSerGlyValVal 151
           :||:::||||:::||||||||||||||||||
Db   554 GCAGGTGATGTATGTTATGCTGAGGATGGC-------------ACTGGTGTCGTG 604

Qy   152 AspTyrSerAsnTyrAspMetLysTyrAlaIleArgLeuAspAlaThrGluPhe 171
           ::: |||:::||||||||::
Db   605 GAGTTGTACGGAAAGAGATATGACCTATGCAGTTCGAAAACTGGATAACACTAAGTT 664

Qy   172 Arg---AsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal 190
           ||| :::|||||||||||||||
Db   665 AGATCTCATGAGGGAGAAACTGCCTACATCCGGGTTAAGTTGATGGGCCAGAAGTCCA 724

Qy   191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgGlyProSer 210
           ||| ::::||||||
Db   725 AGTTATGGAAGATCTGATCTCGAAGC---CGTAGTCGTAGCAGAAGCCGTAGCAGAAGC 781

Qy   211 CysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIle---SerProArg 229
           |||||::::||:||||||||
Db   782 -----------AACAGCAGGAGTCGCGACTTACTCCCCAAGGAGGAGAGCAGGAGGATCACCACGC 832

Qy   230 SerArgProLeu---SerArgSerArgSer 238
           |||
Db   833 TATTCTCCCGTCATAGCAGATCTCGCTCT 862

RESULT 14
ABL29431
ID   ABL29431 standard; DNA; 1478 BP.
XX
AC   ABL29431;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 39766.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.

```
Db  540 GTTGAGTTCCTGCGCCACGAGGACATGAAGTACGCAATCAAAAAATTGGACGACTCTCGC 599

Qy  171 PheArgAsnAlaPheSerSer---AlaTyrIleArgValArgGluTyrGluSerArgSer 189
            ||||||:::                          :::|||
Db  600 TTCCGATTCGCATGAGGGCGGAGGTTGCCTACATTCGCGTACGCGAG---GATAGCGGTGAT 656

Qy  190 ValSerArg-------------------------------------------SerProAsp 195
            ||||||
Db  657 AACGACAGAGACGGTGGTGGTGGCGGCCAGCGTGGCGGTGGTGGCCGGCGCGGAGGT 716

Qy  196 AspSerLysSerTyrArgSerArgGlyProSerCysSerTyrSerSer 215
            |||||||| ||||||    ||||:::|||
Db  717 GGAAGCCGTGACTACCGGCACAGATCCCGTTCGCGC-----TCATTCTG--- 761

Qy  216 LysSerArgSerValSerProAlaArgSerIleSerProArgSerProLeuSerArg 235
        ------------------------------------------TCGAGGCCGCGCCGTCGT 779
Db  762 ------

Qy  236 SerArgSerLeuTyrSerSerValSerArgGlySerLeuLeuArgAlaGlyAspTrp 255
            ||||||
Db  780 GGCACTCCCACATACTCCCCAGTGCGACGT-------------------------- 809

Qy  256 IleSerGlnSerArgLysSerArgSerArgSerArgSerAsnProValSerPro 275
            :::|||    |||:::
Db  810 --------CAATCTTATTCCAGTCTCGCTCACGTTCCACTATTAAACACAACCA 857

Qy  276 ValIle 277
Db  858 AGCATT 863
```

```
RESULT 15
ABI99368
ID  ABI99368 standard; cDNA; 1584 BP.
XX
AC  ABI99368;
XX
DT  07-MAR-2002  (first entry)
XX
DE  Mouse ischaemic condition related cDNA sequence SEQ ID NO:264.
XX
KW  Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW  vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS  Mus musculus.
XX
PN  WO200188188-A2.
XX
PD  22-NOV-2001.
XX
PF  18-MAY-2001; 2001WO-JP004192.
XX
PR  18-MAY-2000; 2000JP-0145977.
XX
PA  (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI  Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR  WPI; 2002-034733/04.
XX
PT  Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT  expression levels of particular genes defined in the specification or by
PT  determining the expression profile of a gene group comprising these
PT  genes.
XX
PS  Claim 2; Page 745; 2690pp; English.
XX
CC  The present invention describes a method for examining ischaemic
CC  conditions, comprising measuring the expression levels of particular
CC  genes (I) in a test sample or determining the expression profile of a
CC  gene group in the sample comprising genes selected from (I). The method
CC  is useful for examining the ischaemic condition (e.g. compressive
CC  ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC  expression levels of particular genes (ABI99202 to ABI99912, encoding the
```

```
XX
PR  23-MAR-2000; 2000US-0191637P.
PR  11-JUL-2000; 2000US-00614150.
XX
PA  (PEKE ) PE CORP NY.
XX
Qy  Venter JC, Adams M, Li PWD, Myers EW;
XX
DR  WPI; 2001-655860/75.
XX
PT  New isolated nucleic acid detection reagent for detecting 1000 or more
PT  genes from Drosophila and for elucidating cell signaling and cell-cell
PT  interactions.
XX
PS  Claim 1; SEQ ID NO 39766; 21pp + Sequence Listing; English.
XX
CC  The invention relates to an isolated nucleic acid detection reagent
CC  capable of detecting 1000 or more genes from Drosophila. The invention is
CC  useful in developmental biology and in elucidating cell signalling and
CC  cell-cell interactions in higher eukaryotes for the development of
CC  insecticides, therapeutics and pharmaceutical drugs. The invention
CC  discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC  sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC  ABB72072). The sequence data for this patent did not form part of the
CC  printed specification, but was obtained in electronic format directly
CC  from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 1478 BP; 416 A; 353 C; 347 G; 362 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                                Length:        1478
Score:             1.41e-32               Matches:        136
Percent Similarity: 54.64%               Conservative:    29
Best Local Similarity: 45.03%            Mismatches:      71
Query Match:        37.90%               Indels:          66
DB:                 4                    Gaps:            10
```

```
US-10-014-927-19 (1-279) x ABL29431 (1-1478)

Qy    1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
          || |||
Db   81 ATGGGATCACGCAACGAGTGCCGTATATATGTGGCAACCTTCCCCAGATATACGCACC 140

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspIleLysIle 40
          ||||||:::
Db  141 AAGGATATCCAGGACCTGTTCCACAAGTTCGGCAAAGTAAACTTTCGTCGATCTGAAAAAT 200

Qy   41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
          ||||| :::|||||
Db  201 CGGCGTGGGCGCCA---TTTGCTTTGTTGGTGAGTTCGAAGAGTCGCGGGATGCCGACGAT 257

Qy   61 AlaIleTyrGlyArgAspGlyTyrArgAspPheAspGlyCysLysArgGluArgIleAla 80
          ||||||
Db  258 GCCGGTGAAGGCGGCGCGACGCGCTCACGATGGGTATCGTCTGCGCCGTGGAGTTCCCG 317

Qy   81 HisGlyGlyArgAspPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSer 100
          |||||||:::
Db  318 CGGGGCGGT-------GGTTCCTGGAAGCTACCGCGGCGCAACCCGCAATGACCGAAGC 368

Qy  101 Arg-------------------------AlaProSerArgArgAspTyrTyrArg 110
          ||||||:::::||
Db  369 CGCCACGGTGGGGACGGACGGATGGGCGACGCGCCAGCCCAAGCGGCTCGCAGTACCGGC 428

Qy  111 ValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArg 130
          |||||| |||
Db  429 GTCATGGTTACTGGACTGCCCCGCCTCCGACGTGGCCAAGATCCTCAAGGATCACATGCGC 488

Qy  131 LysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyVal 150
          :::|||||||:::|||:::::||
Db  489 GAGGCCGCGACCGTCTGCTTCGCGGACACTTACAAAGATGGT--------TCCGGCGTC 539

Qy  151 ValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLysLeuAspAlaThrGlu 170
          |||::::::  |||||||:::||||||||
```

```
Db  859 ATGGGCCCAGAAGTCCAAGTTATGGAAGATCTGATCTCGAAGC---CGTAGTCGTAGCA 915
        ||||||                                        |||::::|||
Qy  205 rgSerArgGlyProSerCysSerTyrSerArgSerLysSerSerArgSerValSerArgProAlaArgS 225
        |||||||                       ::|||||||||||||||
Db  916 GAAGCCGTAGCAGAAGC--------AACAGCAGGAGTGCGAGTTACTCCCCAAGGAGAA 966
        |||||
Qy  225 erIle--SerProArgSerArgProLeu---SerArgSerArgSer 238
        ||:::||                      |||||||||||||||
Db  967 GCAGAGGATCACCACGCTATTCTCCCGTCATAGCAGATCTCGGCTCT 1013

Search completed: March 26, 2004, 18:26:37
Job time : 505 secs
```

```
CC  protein sequences in ABB57020 to ABB57374) or by determining the
CC  expression profile of a gene group comprising these genes. The expression
CC  levels or expression profiles produced by these genes are used as an
CC  indicator when screening for ischaemic condition-improving drugs or
CC  therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC  primers for a mouse ischaemic condition related sequence, which are used
CC  in the exemplification of the present invention
XX
SQ  Sequence 1584 BP; 374 A; 332 C; 428 G; 450 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:              5.22e-32          Length:       1584
Score:                  543.00            Matches:      139
Percent Similarity:     52.68%            Conservative: 28
Best Local Similarity:  43.85%            Mismatches:   42
Query Match:            37.42%            Indels:       109
DB:                                       Gaps:         9
```

US-10-014-927-19 (1-279) x ABI99368 (1-1584)

```
Qy    9 IleTyrValGlyAsnLeuProGlyArgLysCysGluValGluAspLeuPheTyr 28
        || ::|| | ||||::||: ||||||:|||||:|:||||||||
Db  131 ATCTACGTGGGTAACCTACCTCCCGGATATCCGAACCAAGACATCGAGGACGTGTTTTAC 190

Qy   29 LysTyrGlyProIleValAspIleAspLeuLysIleProProGly-ProProGlyTyrAl 48
        |||||| |||||                  ::::|||
Db  191 AAATACCGGCCGCCATCCCGACATCGACCTGAAGAA-CCGCCGCGGGGACCCGCCCTTCGC 249

Qy   48 aPheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyrTy 68
        |||:::||| || ||||||||||:||||||||:||:::::|||||||||||
Db  250 CTTCGTTGAGTTCGAGGACCCGCGAGACGCGGAAGATGCGGGTACGGTCGCGACGGCTA 309

Qy   68 rAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyArg------------ 84
        |:::|||                     |||||
Db  310 CGACTACGACGGCTACCCGGCTGCGGGTAGAGTTCCCCGAAGCGGCCGGGGACCCGCCG 369

Qy   85 --------------ArgPheSerProSerValAs 91
                      ||||::::
Db  370 AGGCGCGGCGGCGGGGGTGAGGCGGCGGCGGCGGCCCCGAGAGGCCGCTATGGCCCG------ 421

Qy   91 pArgTyrSerSerSerTyrSerArgLeuSerArgAlaProSerArgArgSerAspTyrArgVa 111
        |||||::::                  |||||||||
Db  422 CCGTCCAGGCGGTCCGAGAACAGAGT 447

Qy  111 LeuValThrGlyLeuProSerAlaSerPheGlnAspSerTrpGlnAspHisLeuLysMetArgGly 131
        ::::|||                    ||||:::|||||||||
Db  448 GGTGTCTCCTGGACTGCCTCCGAGTGGAAGCTGGCAGGACTTAAAGGATCACATGGCTGA 507

Qy  131 sAlaGlyAspValCysPheSerGluValPheProAspProArgArgLysGlyMetSerGlyValVa 151
        ||:::|||::|||                     ||||||
Db  508 GGCAGGTGATGATGTTACGCGTGGATGTTTACCGGAGATGGC--------ACTGGGTCGT 558

Qy  151 lAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPh 171
        |||:::::                   |||::|||||||||||
Db  559 GGAGTTTGTACGGAAAGAAGAATGACCTATCGAGTTCGAAAACTGGATAACACTAAGTT 618

Qy  171 eArg------------------------------------------Asn-AlaPheSer- 176
        |||                                           ::::|||
Db  619 TAGATCTCACGAGGTAGGTTATACACTTATTCTTTTTTTTGGCCAGAATTGGATACAGTT 678

Qy  176 ------------------------------------------------ 176
Db  679 TTCTTAACAGTGGAATTGAAGTAAGGATACAGGCAAGGGTGTTCACGTAAATTACCAG 738

Qy  176 ------------------------------------------------ 176
Db  739 AGCCCTGATCTGTCTTTGTATTCGTTCAGCTTGTCTGAAGACGAGGTGAAAGCTTAGATCT 798

Qy  177 ----------------------SerAlaTyrIleArgValArgGluT 185
                              ::::||||| |||||
Db  799 TTCAATGAAAGTTCTGTCTATCCAATAGGGAAACTGCCACATCCGGGTTAAGTTG 858

Qy  185 yrGluSerArgSerValSerArgSerTyrArgSerArgSerA 205
```

OM protein - nucleic search, using frame_plus_p2n model

```
         GenCore version 5.1.6
    Copyright (c) 1993 - 2004 Compugen Ltd.
```

Run on: March 26, 2004, 18:17:24 ; Search time 109 Seconds
                                      (without alignments)
                                      1420.471 Million cell updates/sec

Title: US-10-014-927-19
Perfect score: 1451
Sequence: 1 MSSRWNRTIYVGNLPGDIRK.........RSKSRSRSRSNSPVSPVISG 279

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext   0.5
               Ygapop 10.0 , Ygapext   0.5
               Fgapop 6.0 , Fgapext   7.0
               Delop   6.0 , Delext   7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

```
Command line parameters: -DEV=xlh
-MODEL=frame+_p2n.model
-Q=/cgn2_1/USPTO_spool/US10014927/runat_24032004_152926_9928/app_query.fasta_1.455
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITs=Bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=10 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10014927 @CGN 1.69 @runat_24032004_152926_9928 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :   Issued_Patents_NA:*

```
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	541.5	37.3	5285	4	US-09-402-328-1	Sequence 1, Appli
2	418.5	28.8	1534	4	US-09-976-594-961	Sequence 961, App
3	254.5	17.5	284	4	US-09-313-294A-5971	Sequence 5971, Ap
4	214.5	14.8	2886	4	US-09-976-594-960	Sequence 960, App
5	208	14.3	955	4	US-09-620-312D-524	Sequence 524, App
6	187.5	12.9	193	4	US-09-702-705-1719	Sequence 1719, Ap
7	187.5	12.9	193	4	US-09-736-457-1719	Sequence 1719, Ap
8	187.5	12.9	193	4	US-09-671-325-1719	Sequence 1719, Ap
9	179	12.3	1015	4	US-09-023-655-867	Sequence 867, App
10	165	11.4	479	4	US-09-621-976-3128	Sequence 3128, Ap
11	160	11.0	1894	2	US-08-935-450-7	Sequence 7, Appli
12	160	11.0	1894	4	US-09-338-123-7	Sequence 7, Appli
13	154.5	10.6	1278	3	US-09-613-182-14	Sequence 14, Appl
14	154.5	10.6	2846	3	US-09-613-182-5	Sequence 5, Appli
15	146.5	10.1	488	4	US-09-621-976-2159	Sequence 2159, Ap
16	146	10.1	6775	4	US-09-620-312D-289	Sequence 289, App
17	143.5	9.9	320	4	US-09-313-294A-3881	Sequence 3881, Ap
18	139	9.6	3283	3	US-09-061-709-8	Sequence 8, Appli
19	139	9.6	3283	3	US-09-899-651-8	Sequence 8, Appli
20	139	9.6	3412	3	US-09-061-709-6	Sequence 6, Appli
21	139	9.6	3412	3	US-09-899-651-6	Sequence 6, Appli
22	137	9.4	2228	1	US-07-726-607C-1	Sequence 1, Appli
23	137	9.4	2228	1	US-07-843-949A-1	Sequence 1, Appli
24	137	9.4	2228	2	US-08-218-978-1	Sequence 1, Appli
c 25	136.5	9.4	4248	4	US-10-164-595-53	Sequence 53, Appl
26	133	9.2	3034	4	US-09-976-594-636	Sequence 636, App
27	132.5	9.1	1740	4	US-09-643-597-347	Sequence 347, App
28	132.5	9.1	1740	4	US-09-542-615A-347	Sequence 347, App
29	132.5	9.1	1740	4	US-09-606-421B-347	Sequence 347, App
30	131.5	9.1	4159	3	US-09-061-709-4	Sequence 4, Appli
31	131.5	9.1	4159	4	US-09-899-651-4	Sequence 4, Appli
32	131.5	9.1	4181	4	US-09-643-597-175	Sequence 175, App
33	131.5	9.1	4181	4	US-09-480-884A-175	Sequence 175, App
34	131.5	9.1	4181	4	US-09-542-615A-175	Sequence 175, App
35	131.5	9.1	4181	4	US-09-606-421B-175	Sequence 175, App
36	130.5	9.0	781	4	US-09-370-838-150	Sequence 150, App
37	130.5	9.0	1401	1	US-07-843-949A-3	Sequence 3, Appli
38	130.5	9.0	1401	2	US-08-218-978-3	Sequence 3, Appli
39	130.5	9.0	1477	4	US-09-976-594-794	Sequence 794, App
40	128	8.8	522	4	US-09-621-976-3652	Sequence 3652, Ap
41	127.5	8.8	547	4	US-09-370-838-168	Sequence 168, App
42	127	8.8	582	3	US-09-385-982-421	Sequence 421, App
43	126.5	8.7	548	4	US-09-370-838-174	Sequence 174, App
44	126	8.7	1467	1	US-07-881-075-50	Sequence 50, Appl
45	126	8.7	1467	1	US-08-120-827-50	Sequence 50, Appl

ALIGNMENTS

```
RESULT 1
US-09-402-328-1/c
; Sequence 1, Application US/09402328
; Patent No. 6365728
;
; GENERAL INFORMATION:
;  APPLICANT: Purdue Research Foundation,
;             Hodges, Thomas K.
;             Lysnik, Leszek A
;  TITLE OF INVENTION: Regulatory Element For Expressing Genes
;                      In Plants
;  NUMBER OF SEQUENCES: 6
;  CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Barnes & Thornburg
;     STREET: 11 S. Meridian
;     CITY: Indianapolis
;     STATE: Indiana
;     COUNTRY: USA
;     ZIP: 46204
;  COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/402,328
;     FILING DATE: 05-No. 6365728-1999
;     CLASSIFICATION: <Unknown>
;  ATTORNEY/AGENT INFORMATION:
;     NAME: Breen, John P.
;     REGISTRATION NUMBER: 38,833
;     REFERENCE/DOCKET NUMBER: 3220-29933
;  TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (317) 231-7745
;     TELEFAX: (317) 231-7433
; INFORMATION FOR SEQ ID NO: 1:
;    SEQUENCE CHARACTERISTICS:
```

```
          LENGTH: 5285 base pairs
          TYPE: nucleic acid
          STRANDEDNESS: double
          TOPOLOGY: linear
  MOLECULE TYPE: DNA (genomic)
  HYPOTHETICAL: NO
  ANTI-SENSE: NO
  ORIGINAL SOURCE:
     ORGANISM: Arabidopsis thaliana
  SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-402-328-1

Alignment Scores:
Pred. No.:
Score:                    541.50        Length:         5285
Percent Similarity:       35.31%        Matches:         180
Best Local Similarity:    30.56%        Conservative:     28
Query Match:              37.32%        Mismatches:       44
DB:                            4        Indels:          338
                                        Gaps:             14

US-10-014-927-19 (1-279) x US-09-402-328-1 (1-5285)

Qy     1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys  20
           |||||  ||||||  |    ||   ::||  |||   ||||||||||||:::
Db  3693 ATGAGCAGCCGGTCGAGTAGAACGATTTACGTCGGGGACCCTTCCCCGGCGATATCCGTGAA 3634

Qy    21 CysGluValGluAspLeu--------------------------------------------  26
            |||||||||||||:|
Db  3633 AGAGAAGTTGAAGACTTGTTCAGTAAGGTAAATTACTCGTTTCTCCTAARTTAAGAACTC 3574

Qy    26 ------------------------------------------------------------  26
Db  3573 AGAACCTCAGGATCGAATTTGTGTTATGTGATGAAATGTCTTTACAAGATTTGCATGTG 3514

Qy    26 ------------------------------------------------------------  26
Db  3513 AAAGAAATTGAATGACATCTTCGGAATTGTTGGATTCTGATTTATTGGAAGAAGTTTCC 3454

Qy    27 ----------PheTyrLysTyrGlyProIleValAspIleAspLeu--------------  38
                    |||||||||||||||||||||||||||||||
Db  3453 TTAAGAATTGTTCTTCCCTATTTTTTTTG-AGTAGACCTGTTGTTCAAATCGATTTG 3395

Qy    39 LysIleProProArgProProGlyTyrAlaPheValGlu---------------------  51
            ||||||||||||||||||||||||||||
Db  3394 AAGATTCCGCCCAGGCCTCCAGGCTATGCATTCGTCGAGGTCAGATTTTTATGCAACCA 3335

Qy    51 ------------------------------------------------------------  51
Db  3334 CATGTTTTATAAAGGCTTGTGTGTTGATTGATTTACAGTTAGGCTTAGGTGTTGTTGAGTGT 3275

Qy    51 ------------------------------------------------------------  51
Db  3274 TCATGAGTTTCAAATGTTGGGAGATACTACCACTAATGTTTTGTGATATATGTGATTTA 3215

Qy    52 -PheGluAspProArgAspAlaAspAlaIleTyrGlyAspGlyTyrAspPheAs-----  71
            ||||||||||||||||||||||||||||||||||
Db  3214 GTTTGAGGATGCTCGTGATGCTGATGCGGTCAATTTATGGCCGTGATGGTTATGACTTTGA 3155

Qy    71 pGlyCysArgLeuArg------------------------------------------  76
            ||||||
Db  3154 TGGGCATCATTACGGGTTTGTTGTTAATCATCAAATTTTACACAAACTTTGTA 3095

Qy    77 ----------------ValGluIleAlaH-----------------------------  81
                           |||::|||
Db  3094 TTGTTGACTGGTAGCCAAACATATATTTATGTGTTGGCCATTGCAGTGGAACTAGCTC 3035

Qy    81 isGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAla-----  99
            |  |||||||||||||||||||||||||||||
Db  3034 ATGGTGGGAGGCGGTTCATCACCATCGATGCA--------CGCGGTAGTTATAGTGGTCGTG 2984

Qy   100 ----------------------------SerArgAlaProSerArgArgS-------- 107
                                         |||||||||||
Db  2983 GTCGTGGCCGTCGTCGTGGTGGTCGTGGTGGTGGTCGTCGTCGTGGCCATCTAGGAGAT 2924
```

```
Qy   107 erAspTyrArg-----------------------------------------------  110
            |:::||||
Db  2923 CAGAGTACCGCCGCGTACATGATGATGTTTAGTGTATTCATAGTTTAGGAGAA 2864

Qy   111 --------------------------------ValLeuValThrGlyLeuProProSerAlaS  121
                                          |||:|||||||||||||||||||
Db  2863 ATCACTAAATTAAGACATTTCCACTATACAGTTGTAGTGTCAGTTTGCCTTCATCTGGT 2804

Qy   121 erTrpGlnAsp----------------------------------------------  124
            ||||||||||
Db  2803 CCTTGGCAAGACCCTCAAGGTGATTACACAGTTGTGGACAGATTCCTCATTGTCGTTCCTAT 2744

Qy   125 --------------------------------LeuLysAspHisMetArgGlys  131
                                          |||||||||||
Db  2743 ATTTTCTACGATTCTGAATGTTGACATATCACTTGTTTGTTTTGAAGGATCACATGCGTAAA 2684

Qy   132 AlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer---------  148
            ::||||||||||||||||||||||||::||
Db  2683 GGAGGAGAAGTTGTTTTTCTCAAGTGTTTCGTGATGGTAGAGTAAAATCCTCATGTCTT 2624

Qy   148 -----------------------------------------------------------  148
Db  2623 ACCCAAGAATAGAGATGTGTCTAAAGCTGTTTGGGTTTCTGCCTTTAAATCCTGTTTTA 2564

Qy   148 -----------------------------------------------------------  148
Db  2563 ACTAGTCTGTGCATGCTTTAAAAAATCATATCGTGGACTATTTATTTACAGTGACTACAACT 2504

Qy   149 GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArg------------  164
            |||||||||||||||||||||||:::|||
Db  2503 GGAATTGTAGATTATACCAGCTACCGAGAACATGAAATATGCGGTGAGATGTTTTTCTCC 2444

Qy   164 -----------------------------------------------------------  164
Db  2443 TCTTCATATTATTTGATTTTCTGTGTGGGACTGTGGCTGAGAAGTGTGTTCATCCTGATT 2384

Qy   165 --------------------------------LysLeuAspAlaThrGluPheArgAr  172
                                          |||:||||||||
Db  2383 ACTCATGATTATATTCTTTATTTGTTCCAGATAAAAAAAAGCTCGATGACACAGAGTTTCG 2324

Qy   172 gAsnAlaPheSerSerAlaTyrIleArg-------------------------------  181
            ||||||||||||||:::|||
Db  2323 GAATGCGTTTTCTCATGAATATGTTCGGGTATGTTAAAGACAAGAATTCTGAGAGCTGCA 2264

Qy   182 -----------------------------------------------------------  182
                                                      ValArgGluTyrGluSerA  188
Db  2263 TTCGTTTTATTGCACTTAACATTTCCCATGTCATTACAGGTTAGAGAATATGATTCAA 2204

Qy   188 rgSerValSerArgSerProAspAspSerLysSerTyr---ArgSerArgSerSerA  207
            |||||||||      :::|||
Db  2203 GGAGGGATTCCAGGAGCCCCAGCCCGGGAAGATCCTATTCTAAGGACCCGCAGCCGTGGAC 2144

Qy   207 rgGlyProSerCysSerSerTyrSerSerArgSerValSerProAlaSerProAlaArgSerIleS  227
            ||||||||||          ::||||
Db  2143 GGAGTCCCAGC----------CGTAGTCGTAGCCGAGCCAGCGGAGCAAGAGCAGAGAA 2093

Qy   227 erProArgSerArgProLeuSerArgSerLeuTyrSerArgSerLeuTyrSerArgSerG  247
            ||||||||::::      :::
Db  2092 GTCCAAAGGCTAAATCTTTGGTAGATCGGCCTGCA-----AAATCTACACATCGAGATCT- 2040

Qy   247 lySerLeuLeuArgAlaGlyAlaAspTrpIleSerGlnSerArgSerLysSerArgS  267
            |||:::|||          :::
Db  2039 --------------------------------CCTCGCTCTCCGCTCCCGCTCTAAGT 2015

Qy   267 erArgSerAsnSerPro  272
            ||||||||
Db  2014 CGAGGTCACTGTCTCCA 1998
```

RESULT 2
US-09-976-594-961
; Sequence 961, Application US/09976594
; Patent No. 6673549

```
; GENERAL INFORMATION:
;   APPLICANT: Furness, Michael
;   APPLICANT: Buchbinder, Jenny
;   TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;   FILE REFERENCE: PA-0041 US
;   CURRENT APPLICATION NUMBER: US/09/976,594
;   CURRENT FILING DATE: 2001-10-12
;   PRIOR APPLICATION NUMBER: 60/240,409
;   PRIOR FILING DATE: 2000-10-12
;   NUMBER OF SEQ ID NOS: 1143
;   SOFTWARE: PERL Program
; SEQ ID NO 961
;   LENGTH: 1534
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   OTHER INFORMATION: Incyte ID No. 6673549 988704.26
US-09-976-594-961

Alignment Scores:
Pred. No.:              6.33e-29
Score:                  418.50
Percent Similarity:     52.63%
Best Local Similarity:  42.46%
Query Match:            28.84%
DB:                     4

US-10-014-927-19 (1-279) x US-09-976-594-961 (1-1534)           Length:         1534
                                                                Matches:         121
                                                                Conservative:     29
                                                                Mismatches:       90
                                                                Indels:           45
                                                                Gaps:             10

Qy    9  IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr  28
Db  188  GTATTCATCGGGAGACTAAATCCAGCGGCCAGGGAGAAGGACGTGGAAGAATCTTCAAG  247

Qy   29  LysTyrGlyProIleValIleAspIleArgLysIleProGluProProArgProGlyTyrAla  48
Db  248  GGATATGGACGCGGATAAGAGATATTGATCTGAAA----------------AGAGGCTTTGGT  292

Qy   49  PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr  68
Db  293  TTTGTGGAATTTGAGGATCCAAGGGATGCAGATCGTGTGTATGAGCTTGATGGAAAA  352

Qy   69  AspPheArgGlyCysArgLeuArgValGluValGluIleAlaHis------------Gly  82
Db  353  GAACTC------TGTAGTGAAAGGGTTACTATTGAACATGCTAGGGCTCGGTCACGAGGT  406

Qy   83  GlyArgArgPheSerProSerValAspArgTyrSerSerSerAlaSerArg---  101
Db  407  GGAGA---------GGTAGGAGGACGATACTCCTGACCGTTTAGTAGTCCCAGACCT  454

Qy  102  ------------------AlaProSerArgArgSerAspTyrArgSerValLeuValThrGly  115
Db  455  CGAAATGATAGAGCAAAATGTCCACCTGTAAGAACAGAAAATCGTCTTATAGTTGAGAAT  514

Db  116  LeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspVal  135
Db  515  TTATCCTCAAGAGTCAGCTGGCAGGATCCAAAGATTTCATGAGACAAGCTGGGGAGTA  574

Qy  136  CysPheSerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsn  155
Db  575  ACGTTTGCGGATGCACGCGACCACCGACCTAAATTAAAT--GAAGGGGTGGTTGAGTTTGCCTCT  631

Qy  156  TyrAspAspMetLysTyrAlaIleIleArgLysLeuAspAlaThrGluPheArgAsnAlaPhe  175
Db  632  TATGTGACTTAAGAAATGCTATTGAAAAACTTTCTGAAAGGAAATA--AATGGGAGA  688

Qy  176  SerSerAlaTyrIleArgValArgGluTyrGluSerValSerArgSerProAsp  195
Db  689  AAAATAAAATAATTGAAGAGCAGCAAAAGGCACAGTAGTCAAGAAGCAGGTCTCGATCC  748

Qy  196  AspSerLysSerTyrArgSerArgSerArgArgSerArgGlyProSerCysSerSer  215
Db  749  CGGACCAGAAGT---TCCTCTAGGTCTCGGTTCGTCGTAGCCGATCCCGTTCCCGTAGTCGCAAATCT  805

Qy  216  LysSerArgSerValSerProAlaArgSerIleSerSerProArgProArgProLeuSerArg  235
Db  806  TACAGCCGGTCA------------------------------AGAAGCAGGAGCAGGAGCCGG  838

Qy  236  SerArgSerLeuTyrSerSerValSerArgSer-----------  246
Db  839  AGCCGGAGCAAGTCCCGTTCTGTAGTAGGTCTCCCGTGCCTCGAGAAGAGCCAGAAACGT  898

Qy  247  GlySerLeuLeuArgAlaGlyAlaAspTrpIleSerArgSerGlnSerArgLysSerArgArg  266
Db  899  GGTTCTTCAAGTAGATCTAAGGTCCAGCATCTGTGGATCGGCAGAGGTCCGGTCCCGA  958

Qy  267  SerArgSerAsnSer  271
Db  959  TCAAGGTCCAGATCA  973


RESULT 3
US-09-313-294A-5971
; Sequence 5971, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;   APPLICANT: Lalgudi, Raghunath V.
;   APPLICANT: Ito, Laura Y.
;   APPLICANT: Sherman, Bradley K.
;   TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;   FILE REFERENCE: PL-0017 US
;   CURRENT APPLICATION NUMBER: US/09/313,294A
;   CURRENT FILING DATE: 1999-05-14
;   NUMBER OF SEQ ID NOS: 7600
;   SOFTWARE: PERL Program
; SEQ ID NO 5971
;   LENGTH: 284
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc feature
;   OTHER INFORMATION: Incyte ID No. 6476212  700350992H1
;   NAME/KEY: unsure
;   LOCATION: 70, 83, 238, 258-259, 283
;   OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5971

Alignment Scores:
Pred. No.:              5.12e-15          Length:         284
Score:                  254.50            Matches:         60
Percent Similarity:     71.58%            Conservative:     8
Best Local Similarity:  63.16%            Mismatches:      21
Query Match:            17.54%            Indels:           6
DB:                     4                 Gaps:             1

US-10-014-927-19 (1-279) x US-09-313-294A-5971 (1-284)

Qy   39  LysIleProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAla  58
Db    3  AAGTCCCCCCAAGACCACCTGGTTATGCTTTTGTTGAGTTTGAAGATCCTCGTGATGCT  62

Qy   59  AspAspAlaIleTyrGlyArgAspGlyTyrAspPheArgGlyCysArgLeuArgValGlu  78
Db   63  GAGGAGGAGAATTGCTGAACGNGATGAATCACAACTTTGATGGACACCGGTCTAAGAGTGGAG  122

Qy   79  IleAlaHisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSer  98
Db  123  GCTGCTCATGGTGGTAGGAGGAATGCTTCCTCGCCATGATCGTTCAAGTGGCTTGGTGTGGCG  182

Qy   99  Ala--------SerArgAlaProSerArgArgSer--AspTyrArgValLeuVal-ThrG  115
Db  183  GTGGTGGAGCACGTACGTGTGTCGAGACACTCCAGACATGATCGTGTTCTTGTCNACTG  242

Qy  115  lyLeuProProSer-AlaSerTrpGlnGlnProGlnLeuLysAsp  127
Db  243  GAATGCCTTCTTCTGNNATCAGGGCAGGATTTAAAGGAT  281
```

RESULT 4
US-09-976-594-960
; Sequence 960, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 960
; LENGTH: 2886
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 988704.25
; NAME/KEY: unsure
; LOCATION: 17, 2848
; OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-960

Alignment Scores:
Pred. No.:
Score: 6.04e-10
Percent Similarity: 53.16%
Best Local Similarity: 43.04%
Query Match: 14.78%
DB:

Length: 2886
Matches: 68
Conservative: 16
Mismatches: 53
Indels: 21
Gaps: 5

US-10-014-927-19 (1-279) x US-09-976-594-960 (1-2886)

```
Qy   123  GlnAspLeuLysAspHisMetArgLysAlaGlyLysAspValCysPheSerGluValPhePro  142
              :::||||||||||||     |||  || ||||| ||:|:|:
Db  2108  AAGGATCTCAAAGATTTCATGAGACAAGCTGGGGATGCACACCGA  2167

Qy   143  AspArgArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAla  162
              ::|    |||||  |||   |||||||| |:|||
Db  2168  CCTAAATTAAAT---GAAGGGGGTGGTTGAGTTTGCCCTTTATGGTGACTTAAAGAATGCT  2224

Qy   163  IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal  182
              
Db  2225  ATTGAAAAACTTTCTGGAAAAGGAAATA---AATGGGAGAAAATAAAATTAATTGAAGGC  2281

Qy   183  ArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer  202
              :::|||  |||||
Db  2282  AGCAAAAGGCACAGTAGGTCAAGAAGCAGGTCCGATCCCGGACCCAGAAGT---TCCTCT  2338

Qy   203  ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerValSerPro  222
              |||||||||
Db  2339  AGGTCTCGTAGCCGATCCCGTTCCCGTAGTCGCAAATTTACAGCCGGTCA--------  2389

Qy   223  AlaArgSerIleSerProArgSerArgSerArgSerLeuSerArgSerLeuTyrSerSer  242
              ||||||||
Db  2390  ------------------AGAAGCAGGAGCAGCAGGAGCCGGAGCCAAGTCCCGTTCT  2431

Qy   243  ValSerArgSer----------------GlySerLeuLeuArgAlaGly  253
              |||||||||
Db  2432  GTTAGTAGGTCTCCCGTGCCTCCCGTCCGTGAGAGCCCAGAAGACCGAGAAGATCTAAG  2491

Qy   254  AspTrpIleSerGlnSerArgSerLysSerArgSerArgSerAsnSer  271
              ::|||||
Db  2492  TCTCCAGCACATCTGTGGATGCGCCAGAGTCCCGGTCCCGATCCAGATCA  2545
```

RESULT 5
US-09-620-312D-524
; Sequence 524, Application US/09620312D
; Patent No. 6569662

; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 524
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(866)
US-09-620-312D-524

Alignment Scores:
Pred. No.:
Score: 4.93e-10
Percent Similarity: 208.00
Best Local Similarity: 43.69%
Query Match: 28.67%
DB: 14.33%

Length: 955
Matches: 84
Conservative: 44
Mismatches: 91
Indels: 74
Gaps: 12

US-10-014-927-19 (1-279) x US-09-620-312D-524 (1-955)

```
Qy     6  AsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAsp  25
              ||||||  ||:||:    :::::
Db   105  AACACCTCCCTGTTCATCAGGAACGTCGCGGACGCCACCAGGCTGGCCGT  164

Qy    26  LeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle--------ProPro  42
              |||     :: ::
Db   165  GAGTTTGGTCGATATGGCCCCATAGTAGACGTTTACATTCCACTTCTACACTCGC  224

Qy    43  ArgProProGlyTyrAlaPheValGluAspProArgAspAlaAspAspAlaIle  62
              |||
Db   225  CGCCCAAGAGGATTTGCTTATGTTCAATTTGAAGATGTTCGAGATGCTGAAGATGCTCTT  284

Qy    63  TyrGlyArgAspGlyTyrArgAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly  82
              ::|
Db   285  TATAACCTCAATAGAAAGTGGTATGTGGCCGTCAGATTGAAATACAGTTTGCACAAGGT  344

Qy    83  GlyArgArg----------------------------PheSerProSerVal  90
              |||:::
Db   345  GATCGCAAAACCACCAGCCAAATGAAATCAAAAGAACGTCATCCTTGTTCTCCAAGTGAT  404

Qy    91  AspArgTyrSerSerTyrSerAlaSerArgAlaProSerArgSerAspTyrArg  110
              |||||||
Db   405  CACAGGAGATCAAGAAGCCCCAGCCCAAAGAGAGAACTCGAAGTAGA----------  449

Qy   111  ValLeuValThrGlyLeuProProSerAlaSerTrp--------GlnAspLeuLysAsp  127
```

Db   450 ---------------------------AGTTCTTCATGGGAAGAAATAGAGGCGGTCAGAC 485
                                      |||::::||||||| |||
Qy   128 HisMetArgAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMet 147
                ::::::::::| ||||||
Db   486 AGCCTTAAAGAGTCTCGACACAGGCGGATTTCT----------------------- 518

Qy   148 SerGlyValValAspTyrSerAsnTyrArgAspMetLysTyrAlaIleArgLysLeuAsp 167
                                                   ||||| ||
Db   519 ----------------------TATAGCCAGTCTAAATCTCGTTCCAAATCATTACCA 554

Qy   168 AlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal------ArgGluTyr 185
                                || ::: |||||| |||
Db   555 ----------AGGCGGTCTACCTCAGCAAGCGAGTCAAGAACTCCAAGAAGGAATTT 602

Qy   186 GluSerArgSerValSerArgSerProAspAspSerLysSerArgSerTyrArgSerArgSerArg 205
                ||||||| ::: |||||||
Db   603 GGCTCTAGAGACGGTCAAG-----------TCCAAGTCCTTACAAAGAGAGTCCAAG 650

Qy   206 SerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSer 225
                ||||| ::: |||::||
Db   651 TCAATAGGA-----------------AAATCACAGTCAAGTTCACCTCAAAAGCAG 689

Qy   226 IleSerProArgProLeuSerArgSerArgSerLeuTyrSer--SerValSer 244
                   ::: ||||||| ||||||||||
Db   690 ACTAGCTCAGGAACAAA------TCAAGATCACATGGAAGACATTCTGACTCAATAGCA 743

Qy   245 ArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArg---------- 260
                ||||| |||::::::::|
Db   744 AGATCCCCGTGTAAATCTCCCAAAGGGTATACCAATTCTGAAACTAAAGTACAAACAGCA 803

Qy   261 -----SerLysSerArgSerArgSerArgSerAsnSer 271
                    ||||| ||||||||
Db   804 AAGCATTCTCATTTCGGTCACCATTCCAGATCTCGAAGT 842

RESULT 6
US-09-702-705-1719
; Sequence 1719, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   APPLICANT: Fan, Liqun
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C14
;   CURRENT APPLICATION NUMBER: US/09/702,705
;   CURRENT FILING DATE: 2000-10-30
;   NUMBER OF SEQ ID NOS: 1833
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1719
;   LENGTH: 193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-702-705-1719

Alignment Scores:
Pred. No.:   Length: 193
Score: 3.78e-09   Matches: 40
Percent Similarity: 187.50   Conservative: 8
Best Local Similarity: 69.57%   Mismatches: 12
Query Match: 57.97%   Indels: 9
DB: 12.92%   Gaps: 2
                     4

US-10-014-927-19 (1-279) x US-09-702-705-1719 (1-193)

Qy   53 GluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly 72

Db   3 GAGGACCCCGAGATGCAGAGAGATGCTATTTATGGAAGAAATGGTTATGATTATGGCCAG 62
            |||||:::::|||||||||||||||||| |||||::::||||||
Qy   73 CysArgLeuArgValGluIle------AlaHisGlyGlyArgArgPheSerProSerVal 90
                 ||||::::||| :::|||||||
Db   63 TGTCGGCTTCGTGTGGAGTCCCCAGGACTTATGGAGGTCGGGGTCGGGGTCGGCCCGTGT 122

Qy   91 AspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgSerArgAspTyrArg 110
                ||| ::: ||||||::::|
Db   123 GGGAGG-------------------------AATGGGCCTCCTACAAGAAGATCTGATTTCCGA 161

Qy   111 ValLeuValThrGlyLeuProProSer 119
                ||||:::::|||::|||||
Db   162 GTTCTTGTTTCAGGACTTCCTCCGTCA 188

RESULT 7
US-09-736-457-1719
; Sequence 1719, Application US/09736457
; Patent No. 6509948
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Lodes, Michael A.
;   APPLICANT: Fanger, Gary
;   APPLICANT: Vedvick, Tom
;   APPLICANT: Carter, Darrick
;   APPLICANT: Retter, Marc
;   APPLICANT: Mannion, Jane
;   APPLICANT: Fan, Liqun
;   APPLICANT: Wang, Aijun
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;   TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.478C15
;   CURRENT APPLICATION NUMBER: US/09/736,457
;   CURRENT FILING DATE: 2000-12-13
;   NUMBER OF SEQ ID NOS: 1864
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1719
;   LENGTH: 193
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-736-457-1719

Alignment Scores:
Pred. No.:   Length: 193
Score: 3.78e-09   Matches: 40
Percent Similarity: 187.50   Conservative: 8
Best Local Similarity: 69.57%   Mismatches: 12
Query Match: 57.97%   Indels: 9
DB: 12.92%   Gaps: 2
                     4

US-10-014-927-19 (1-279) x US-09-736-457-1719 (1-193)

Qy   53 GluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly 72
               |||||:::::|||||||||||||||||| |||||::::||||||
Db   3 GAGGACCCCGAGATGCAGAGAGATGCTATTTATGGAAGAAATGGTTATGATTATGGCCAG 62

Qy   73 CysArgLeuArgValGluIle------AlaHisGlyGlyArgArgPheSerProSerVal 90
                ||||::::||| :::|||||||
Db   3 TGTCGGCTTCGTGTGGAGTCCCCAGGACTTATGGAGGTCGGGGTCGGGGTCGGCCCGTGT 62

Qy   91 AspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgSerArgAspTyrArg 110
                ||| ::: ||||||::::|
Db   63 GGGAGG-------------------------AATGGGCCTCCTACAAGAAGATCTGATTTCCGA 122

Qy   111 ValLeuValThrGlyLeuProProSer 119
                ||||:::::|||::|||||
Db   91 GTTCTTGTTTCAGGACTTCCTCCGTCA 188

RESULT 8
US-09-671-325-1719
; Sequence 1719, Application US/09671325
; Patent No. 6667154

Qy   53 GluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly 72

; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1719
; LENGTH: 193
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-1719

Alignment Scores:
Pred. No.:   3.78e-09   Length: 193
Score: 187.50   Matches: 40
Percent Similarity: 69.57%   Conservative: 8
Best Local Similarity: 57.97%   Mismatches: 12
Query Match: 12.92%   Indels: 9
DB: 4   Gaps: 2

US-10-014-927-19 (1-279) x US-09-671-325-1719 (1-193)

```
Qy   53  GluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly   72
Db    3  GAGGACACCCCGAGATGCAGAGGATGCTATTTATGGAGAAATGGTTATGATTATGGCCAG   62

Qy   73  CysArgLeuArgValGluIle------AlaHisGlyGlyArgPheSerProSerVal      90
Db   63  TGTCGGCTTCGTTGTGTGAGTTCCCAGGACTTATGAGGTCGGGGTCGGCCCCGTGGT    122

Qy   91  AspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgAlaSerAspTyrArg  110
Db  123  GGGAGG------------------AATGGGCCTCCTACAAGAAGAAGATCTGATTTCCGA  161

Qy  111  ValLeuValThrGlyLeuProProSer  119
Db  162  GTTCTTGTTTCAGGACTTCCTCCGTCA  188
```

RESULT 9
US-09-023-655-867
; Sequence 867, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 867:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1015 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1049089
US-09-023-655-867

Alignment Scores:
Pred. No.:   2.36e-07   Length: 1015
Score: 179.00   Matches: 50
Percent Similarity: 53.98%   Conservative: 11
Best Local Similarity: 44.25%   Mismatches: 32
Query Match: 12.34%   Indels: 20
DB: 4   Gaps: 6

US-10-014-927-19 (1-279) x US-09-023-655-867 (1-1015)

```
Qy    9  IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr   28
Db  234  GTATTCATCGGAGACTAAATCCAGCGGCCAGGGAGAAGACGTGGAAGATTCTTCAAG    293

Qy   29  LysTyrGlyTyrProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla   48
Db  294  GGAIATGGACGGATAAGAGATATTGATCTGAAA-----------------AGAGGCTTTGGT  338

Qy   49  PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr   68
Db  339  TTTGTGGAATTTGAGGATCCAAGGGATGCAGATGATGCTGTGTGATGAGCTTGATGGAAAA  398

Qy   69  AspPheAspGlyCysArgLeuArgValGluIleAlaHis-------------Gly   82
Db  399  GAACTC--------TGTAGTGAAAGGGTTACTATTGAACATGCTAGGGCTCGGTCACGAGGT  452

Qy   83  GlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAla  102
Db  453  GGAAGA------------GGTAGAGGACGATACTCTGACCGTTTTAGT---AGTCGCAGA   497

Qy  103  ProSerArgSerAspTyrArgValLeuValThrGly  115
Db  498  CCT------CGAAATGATAGACGGTATGTGAAGGGTGGA  530
```

RESULT 10
US-09-621-976-3128
; Sequence 3128, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.J.Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm

```
; SEQ ID NO 3128
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 165..407
US-09-621-976-3128

Alignment Scores:
Pred. No.:            1.54e-06
Score:                165.00
Percent Similarity:   54.13%
Best Local Similarity: 44.95%
Query Match:          11.37%
DB:                   4

Matches:     479
Conservative: 49
Mismatches:  10
Indels:      29
Gaps:        21
                      6
```

```
US-10-014-927-19 (1-279) x US-09-621-976-3128 (1-479)

Qy    9  IleTyrValGlyAsnLeuProGlyAspIleArgLysGluValGluAspLeuPheTyr   28
           :::::::::|||             :::         :::|||
Db   180  GTATTCATCGGGAGACTAAATCCAGCGGCCAGGGAGGAAGGACGTGGAAAGATTCTTCAAG  239

Qy   29  LysTyrGlyProIleValAspIleAspIleLysLeuLysAspLeuLysIleProAspIleAspLeu-  48
           |||    |||      :::
Db   240  GGATATGGACGGATAAGAGATATTGATCTGAAA----------------AGAGGCTTTGGT  284

Qy   49  PheValGluPheGluAspArgAspAlaAspAlaAspGlyTyrGlyTyrArgGlyAspGlyTyr  68
           |||||||   :::  :::|||||
Db   285  TTTGTGGAATTTGAGGATCCAAGGATGCAGATGATGTGTGTATGAGCTTGATGGAAAA  344

Qy   69  AspPheAspGlyCysArgLeuArgValGluIleAlaHis-------------------Gl  82
           :::
Db   345  GAACTC-----TGTAGTGAAAGGGTTACTACTTGRACAATGCTAGGGCGTCGGTCACGAGGG  398

Qy   82  yGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAl  102
           ||| |||||||                :::||||
Db   399  TGGAAGA---------GGTAGAGGACGATACTCTGACCCGTTTYAGT--AGTCGCAG  443

Qy  102  aProSerArgArgSerAspTyrArg  110
           |||  |||:::
Db  444  ACCT------CGAAATGATAGACGG  462
```

```
RESULT 11
US-08-935-450-7
; Sequence 7, Application US/08935450
; Patent No. 5977311
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; TITLE OF INVENTION: 53BP2 COMPLEXES
; FILE REFERENCE: 7934-054
; CURRENT APPLICATION NUMBER: US/08/935,450
; CURRENT APPLICATION FILING DATE: 1997-09-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)..(1322)
US-08-935-450-7

Alignment Scores:
Pred. No.:            3.08e-05
Score:                160.00
Percent Similarity:   34.13%
Best Local Similarity: 25.40%
Query Match:          11.03%
DB:                   2

Length:      1894
Matches:     96
Conservative: 33
Mismatches:  127
Indels:      123
Gaps:        15
```

```
US-10-014-927-19 (1-279) x US-08-935-450-7 (1-1894)

Qy    9  IleTyrValGlyAsnLeuProGlyAspIleArgLysGluValGluAspLeuPheTyr   28
           :::::::::|||          :::       :::|||         :::||||||
Db   39  CTCTTCATTGGTGGGCTTAATACGAACAATGAGAAAGCTTGAAGCAGTATTTGGC   98

Qy   29  LysIleProProArgLeu------------------LysIleProProArgProPro  45
           |||                                   :::
Db   99  AAATATGGACGAATAGTGGAGTACTCTTGATGAAGACCGTGAAACCAACAAATCAAGA  158

Qy   46  GlyTyrAlaPheValGluPheGluAspProArgAspAlaAspTyrGlyArg  65
           |||:::|||           :::            :::
Db  159  GGATTTGCTTTTGTCACCTTTGAAAGCCCAGCAGCCCAAGGATGCAGCCAGAGACATG  218

Qy   66  AspGlyTyrArgAspPheAspGlyCysArgLeuArgValGluIleAla---------  80
           |||||||         |||      :::||||||
Db  219  AATGGAAAGTCATTAGATGGAAGAACCATCAAGTGGAACAAGCCACCAAACCATCATT  278

Qy   81  HisGlyGlyArgArgPheSerProSerValAspArg---------  92
           ||||||
Db  279  GAAAGTGGTAGACGTGGACCGCCTCCACCTCCAAGAGTAGAGGGTCTT  338

Qy   92                                                  92
Db  339  AGAGTGGAAGAGGAGGAGGAGTGGAGGAACCAGGAGACCTCCCTCACGGGGAGGACACATG  398

Qy   93  -------TyrSerSerSerTyrSerAlaSer---------  100
           :::::::
Db  399  GATGACCGGTGATATTCCATGAATTTAACATGAGTTCTTCCAGGGACCACTCCCAGTA  458

Qy  101  --ArgAlaProSerArgArgSerAspTyrArgValThrGlyLeuPro----  118
           ||||||
Db  459  AAAGAGGACCACCACCAAGAAGTGGG----------GGTCCTCCCTCTCAAG  500

Qy  119  --SerAlaSerTrpGlnAspLeuLysMetArgHisAspHisArgAlaGlyValCysPhe  137
           :::||||||
Db  501  AGATCTGCACCTTCAGGACCCAGTTCCAGTAGCAGTGGAATGGGAGGAGAGCTCCTGTA  560

Qy  138  Ser----------------------------------------GluValPheProAspArgLys  145
           |||
Db  561  TCACGTGGAGAGAGTAGTTATGGAGGTCCACCTCGAAGGGAACCCGCTGCCCTCGTAGA  620

Qy  146  -------------------------GlyMetSerGlyValValAspTyrSerAsnTyr  156
           |||    :::
Db  621  GATGTTTATTGTCTCCAAGATGATGATGGTATTCTACTAAGACAGCTATTCAAGCAGA  680

Qy  157  Asp-----------------------------AspMetLysTyrAlaIle  163
           |||
Db  681  GATTACCCAAGTTCTGTGATACTAGAGATTATGCACCACCAGAGATTATACTTAC  740

Qy  164  ArgLysLeuAspAlaThrGluPheGluArgAsnAlaPheSerSer-----------  177
           |||:::
Db  741  CGTGATTATGTGTCATTCCAGTTCACGTGATGACTATCCATCAAGAGAATATAGGCATAGA  800

Qy  178  --AlaTyrIleArgValArgGluTyrGluSerArgSerProAspAsp  196
           |||:::|||
Db  801  GATGGATATGTCGTGATCGTGACTCATTCAGATCATCCAGTGGAGGTTCCTACAGAT  860

Qy  197  Ser---------LysSerTyr------ArgSerArgSerGlyTyrProSerCysSer  212
           :::||||||              :::|||
Db  861  TCATATGAGAGTTATGGTAACTCACGTAGTGCTCCACCTACACGAGAGGCCCCGCCATCT  920

Qy  213  TyrSerSerLysSerArg---------  218
           |||
Db  921  TATGGTGGAAGCAGTCGCTATGATGATTACAGCAGTCACGTGACGGAATGGTCGGAAGT  980

Qy  219  ------SerValSerProAlaArgSer--------IleSerProArgArgProLeu  233
           :::||||||                       |||
Db  981  CGAGACAGTTACTCAAGCAGCCGAGTGATCTACTCAAGTGGTCGTGATCGGGTTGGC 1040

Qy  234  SerArgSerArgSerLeuTyrSerSerValSerGlySerLeuLeuArgAlaGly  253
           :::
```

```
Db   1041 AGACAAGAAAGAGGGGCTTCCCCCTTCTATGGAAAGG---GGGTACCTCCTCCACGTGAT- 1096

Qy    254 AspTrpIleSerGlnSerArgSerLysSerArgSerArgSerArgSerAsnSer 271
              ::|     |||   :::
Db   1097 TCCTACAGCAGTTCAAGCCGCGCGGACGACCAAGAGGTGGTGGCCGTGGAGGAAGC 1150
```

```
RESULT 12
US-09-338-123-7
; Sequence 7, Application US/09338123
; Patent No. 6627405
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; TITLE OF INVENTION: 53BP2 COMPLEXES
; FILE REFERENCE: 7934-054
; CURRENT APPLICATION NUMBER: US/09/338,123
; CURRENT FILING DATE: 1999-06-22
; EARLIER APPLICATION NUMBER: 08/935,450
; EARLIER FILING DATE: 1997-09-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 1894
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (12)..(1322)
US-09-338-123-7
```

```
Alignment Scores:
Pred. No.:                            1894
Score:              3.08e-05         Matches:             96
                    160.00           Conservative:        33
Percent Similarity:  34.13%          Mismatches:         127
Best Local Similarity: 25.40%        Indels:             123
Query Match:         11.03%          Gaps:                15
DB:                   4
```

```
US-10-014-927-19 (1-279) x US-09-338-123-7 (1-1894)

Qy      9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
          ||||   ::::|||||||    |:|     :::      :::|||
Db     39 CTCTTCATTGGTGGGCTTAATACCGGAAACAAATGAGAAAGCTCTTGAAGCAGTATTGGC 98

Qy     29 LysTyrGlyProIleValAspIleAspLeu--------LysIleProProArgProPro 45
          ||||||:||||   ||    |||           ::|      |||
Db     99 AAATATGAGACGAATAGTGGAAGTACTCTTGATGAAAGACCGTGAAACCAACAAATCAAGA 158

Qy     46 GlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArg 65
          ||:||||:||||||||||:|  |||||| |||||
Db    159 GGATTTGCTTTTGTCACCTTTGAAAGCCCAGCACGCTAAGGATGCCAGCAGCCAGAGACATG 218

Qy     66 AspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla------------ 80
          :::|||       ||||||||
Db    219 AATGGAAAGTCATTAGATGGAAAAGCCATCAAGGTGGAACAAGCCACCAAAACCATCATTT 278

Qy     81 HisGlyGlyArgArgPheSerProSerValAspArg--------------- 92
          ||||||||
Db    279 GAAAGTGGTAGACGTGGACCGCGCCTCCACCTCCAAGAGTAGAGGCCCTCCAAGAGGTCTT 338

Qy     92 --------------------------------- 92
Db    339 AGAGGTGGAAGAGGAGGAAGAGGAGAGTGGGAAGGAAGGCCACCTCCCCTCACGGGGGAGGACCACATG 398

Qy     93 --------TyrSerSerTyrSerAlaSer----- 100
                   :::::::::::|||
Db    399 GATGACGGTGGATATTCCATGAATTTAACATGAGTTCTTCCAGGGGACCACTCCCAGTA 458

Qy    101 ---ArgAlaProSerArgArgSerArgAspTyrArgValLeuValThrGlyLeuProPro--- 118
          ::|||||||
Db    459 AAAAGAGGAGGACCACCACCAAGAAGTGGG--------------GGTCCTCCTCCTAAG 500
```

```
Db   1041 AGACAAGAAAGAGGGGCTTCCCCCTTCTATGGAAAGG---GGGTACCTCCTCCACGTGAT- 1096

Qy    119 ---SerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137
              :::::|
Db    501 AGATCTGCACCTTCAGGACCAGTTCGCAGTGCCAGTAGCCAGTGGAATGGGAGGAGAGCTCCTGTA 560

Qy    138 Ser----------------------GluValPheProAspArgLys 145
          |||                      ||||:::||||:::
Db    561 TCACGTGGAAGAGATAGTTATGGAGGTCCACCTGGAAGGGAACCGCTGCCCTCCTCGTAGA 620

Qy    146 ---------------------GlyMetSerGlyValValAspTyrSerAsnTyr 156
                                |||||||
Db    621 GATGTTTATTTGTTCCAAGAGATGATGGGTATTCTACTAAAGACAGCTATTCAAGCAGA 680

Qy    157 Asp------------------------AspMetLysTyrAlaIle 163
          |||                         |||
Db    681 GATTACCAAGTTCCTGTGATACTAGAGATTATGCACCACCACGAGATTATACTTAC 740

Qy    164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSer------------------- 177
          ||||:::|||                  |:: |||
Db    741 CGTGATTATGGTCATTCCAGTTCACGTGACTATCCATCCAAGAGAATATAGGCGATAGA 800

Qy    178 ---AlaTyrIleArgValArgGluTyrGluSerArgValSerArgSerProAspAsp 196
             |||:::|||||||        ||:::|
Db    801 GATGGAATATGGTCGTGATCGGACTATTCAGATCATCCAAGTGGAGGTTCCTACAGAGAT 860

Qy    197 Ser---LysSerTyr---------ArgSerArgSerArgGlyProSerCysSer 212
          |||      |||||||            :::|||||||
Db    861 TCATATGAGAGTTATGGTAACTCACGTAGTCTCCACCTACACGAGGGCCCGCCATCT 920

Qy    213 TyrSerSerLysSerArg--------------------------------- 218
          ||||||
Db    921 TATGGTGGAAGCAGTCGTGCTATGATGATTACAGCAGCTCACGTGACGGATATGGTGAAGT 980

Qy    219 ------SerValSerProAlaArgSer---------IleSerProArgSerArgProLeu 233
          ||||                     :::|||
Db    981 CGAGACAGTTACTCAAGCAGCCGAAGTGATCTCTACTCAAGTGGTCGTGATCGGGTTGGC 1040

Qy    234 SerArgArgSerLeuTyrSerSerValSerArgSerArgSerLeuLeuArgAlaGly 253
          |||     :::   :::       |||:::|||
Db   1041 AGACAAGAAAGAGGGGCTTCCCCCTTCTATGGAAAGG---GGGTACCTCCTCCACGTGAT- 1096

Qy    254 AspTrpIleSerGlnSerArgSerLysSerArgSerArgSerAsnSer 271
              ::|     :::
Db   1097 TCCTACAGCAGTTCAAGCCGCGCGGACGACCAAGAGGTGGTGGCCGTGGAGGAAGC 1150
```

```
RESULT 13
US-09-613-182-14
; Sequence 14, Application US/09613182
; Patent No. 6294653
; GENERAL INFORMATION:
; APPLICANT: Mayfield, Stephen
; TITLE OF INVENTION: RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT MOLECULES
; FILE REFERENCE: SCR2177S
; CURRENT APPLICATION NUMBER: US/09/613,182
; CURRENT FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: 09/341,550
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: PCT/US98/00840
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 60/035,955
; PRIOR FILING DATE: 1997-01-17
; PRIOR APPLICATION NUMBER: 60/069,400
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 1278
;   TYPE: DNA
;   ORGANISM: Chlamydomonas reinhardtii
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1272)
US-09-613-182-14
```

; Sequence 5, Application US/09613182
; Patent No. 6294653
; GENERAL INFORMATION:
;   APPLICANT: Mayfield, Stephen
;   TITLE OF INVENTION: RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR
;   TITLE OF INVENTION: EXPRESSION OF RECOMBINANT MOLECULES
;   FILE REFERENCE: SCR2177S
;   CURRENT APPLICATION NUMBER: US/09/613,182
;   CURRENT FILING DATE: 2000-07-10
;   PRIOR APPLICATION NUMBER: 09/341,550
;   PRIOR FILING DATE: 1999-07-13
;   PRIOR APPLICATION NUMBER: PCT/US98/00840
;   PRIOR FILING DATE: 1998-01-16
;   PRIOR APPLICATION NUMBER: 60/035,955
;   PRIOR FILING DATE: 1997-01-17
;   PRIOR APPLICATION NUMBER: 60/069,400
;   PRIOR FILING DATE: 1997-12-12
;   NUMBER OF SEQ ID NOS: 14
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
;   LENGTH: 2846
;   TYPE: DNA
;   ORGANISM: Chlamydomonas reinhardtii
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (197)..(2065)
; US-09-613-182-5

```
Alignment Scores:
Pred. No.:
Score:                    154.50          Length:          2846
Percent Similarity:        41.36%         Matches:           67
Best Local Similarity:     22.71%         Conservative:      55
Query Match:               10.65%         Mismatches:       111
DB:                                       Indels:            62
                                          Gaps:              10
                    0.000174

US-10-014-927-19 (1-279) x US-09-613-182-5 (1-2846)

Qy    6 AsnArgThrIleTyrValGlyAsnLeuProGlyValAspIleArgLysCysGluValGluAsp  25
          :::::::      :: ::  ::: ::    ::::::::::   :::
Db  260 AACTCGTCGCTGTACGTCGGTGACCTGCCGTGACCTGGAGAAGGATGTCACCGAGCCCAGCTGTTCGAG 319

Qy   26 LeuPheTyrLysTyrGlyTyrGlyProIleValAspIleAspLeu------LysIleProPro  42
          ||||| |||||||| || ||||||||||  :::
Db  320 CTCTTCTCCTCGGTTGGCCTGCCTGTGCCCGATGCCGTGCCCAGATTCTGTCC            379

Qy   43 ArgProProGlyTyrAlaPheValGluPheGlu------AspProArgAspAlaAsp        59
          ||||||| |:||    :::
Db  380 CGCTCGCTGGGCTACGCCTACGTCAACTACACAGGGCTCGAACCCCCAGGCGTCTGAC       439

Qy   60 AspAlaIleTyrArgGlyTyrArgAspGlyTyrArgAspPheGlyCysArgLeuArgValGluIle 79
          ||||| ||||:||||||||||
Db  440 CGGGCCATGGAGACCCTGACTACCACCATGTCTGAACGCCATCATCGTGG              499

Qy   80 AlaHisGlyGlyArgArgPheSerProValAspArgTyrSerSerTyrSerAla           99
          ::|
Db  500 TCGCAC-------                                                    505

Qy  100 SerArgAlaProSerArgArgSerAsp------TyrArgValLeuValThrGlyLeuPro    117
          ||||||     :: :::
Db  506 ---CGCGACCCTTCGGCCCGCAAGTCGGGCGTCGGCAACATCTTCATCAAGAACCTGGAC   562

Qy  118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe    137
          ::    ||||   :::
Db  563 AAGACCATCACGACGACCCTGCCACGCCTTCGGCCTTCGCAAGATTCTGTCC           622

Qy  138 SerGluValPheProAspArgLysGlyMetSerGlyValValAspArgTyrSerAsnTyrAsp 157
          ||   :::
Db  623 TGCAAGGTTGCCACTGACGGCCAACCGCGTCGAAGGGCTACGGCTTCGTGCACTTCGAG    682

Qy  158 AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgGlyAsnAlaPheSerSer 177
          ::    |||
Db  683 GACCAGGCCCGCTGCCGATCGCCGCCATTCAGACC-------                     715
```

```
Alignment Scores:
Pred. No.:
Score:                    154.50          Length:          1278
Percent Similarity:        41.36%         Matches:           67
Best Local Similarity:     22.71%         Conservative:      55
Query Match:               10.65%         Mismatches:       111
DB:                                       Indels:            62
                                          Gaps:              10
                    5.6e-05

US-10-014-927-19 (1-279) x US-09-613-182-14 (1-1278)

Qy    6 AsnArgThrIleTyrValGlyAsnLeuProGlyValAspIleArgLysCysGluValGluAsp  25
          :::::::      :: ::  ::: ::    ::::::::::   :::
Db  127 AACTCGTCGCTGTACGTCGGTGACCTGCCGTGACCTGGAGAAGGATGTCACCGAGCCCAGCTGTTCGAG 186

Qy   26 LeuPheTyrLysTyrGlyTyrGlyProIleValAspIleAspLeu------LysIleProPro  42
          ||||| |||||||| || ||||||||||  :::
Db  187 CTCTTCTCCTCGGTTGGCCTGCCTGTGCCCGATGCCGTGCCCAGATTCTGTCC            246

Qy   43 ArgProProGlyTyrAlaPheValGluPheGlu------AspProArgAspAlaAsp        59
          ||||||| |:||    :::
Db  247 CGCTCGCTGGGCTACGCCTACGTCAACTACACAGGGCTCGGACCCCCAGGCGTCTGAC       306

Qy   60 AspAlaIleTyrArgGlyTyrArgAspGlyTyrArgAspPheGlyCysArgLeuArgValGluIle 79
          ||||| ||:|||:||||||||||
Db  307 CGCGGCCATGGAGACCCTGAACTACCATGTCTGAACGGCAAGCCTTATGCGCATCATGTGG  366

Qy   80 AlaHisGlyGlyArgArgPheSerProValAspArgTyrSerSerTyrSerAla           99
          ::|
Db  367 TCGCAC-------                                                    372

Qy  100 SerArgAlaProSerArgArgSerAsp------TyrArgValLeuValThrGlyLeuPro    117
          ||||||     :: :::
Db  373 ---CGCGACCCTTCGGCCCGCAAGTCGGGCGTCGGCAACATCTTCATCAAGAACCTGGAC   429

Qy  118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSer 137
          ::    ||||   :::
Db  430 AAGACCATCACGACGACCCTGCCACGCCTTCTCGGCCTTCGCAAGATTCTGTCC          489

Qy  138 SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp    157
          ||   :::
Db  490 TGCAAGGTTGCCACTGACGGCCAACGGCGTCGAAGGGCTACGGCTTCGTGCACTTCGAG    549

Qy  158 AspMetLysTyrAlaIleAlaIleArgLysLeuAspAlaThrGluPheArgGluPheSerSer 177
          ::    |||
Db  550 GACCAGGCCCGCTGCCGATCGCCGCCATTCAGACC-------                     582

Qy  178 AlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSer    197
          :: :::
Db  583 ------GTCAACCAGAGAAGATTGAGGGCAAGATCGTGTACGTGGCCCCCTTCCAGAAG    636

Qy  198 LysSerTyrArgSerArgSerArg--GlyProSerCysSerTyrSerSerLysSe        217
          |:    :: :::
Db  637 CGGCTGACCGCGCCCCAGGGCAAGGACCTTGTACACCAACGTGTTCG-------TC       684

Qy  217 rArgSerValSerProAlaArgSerIleSerProAlaArgSerArg--ProLeuSerArgSe 236
          ::    :::
Db  685 AAGAACTTGCCGCGCGACCATCGGCGACGACGCTGGGCAAGATGGCCACCGAGCACGGC    744

Qy  236 rArgSer------LeuTyrSe                                          241
          |||||
Db  745 GAGAATCACCAGCCAGGCGGGTGGTCATGAGGAGGACGACGACAAGGAGGCTTCGGCTTC   804

Qy  241 rSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSe   261
          |||    :::
Db  805 ATCAACTTCAAGGACGCCGGAGT---CCGCGGCGAAGTGCGTGG------AGTACCTGAAC  855

Qy  261 rLysSerArgSerArgSerArgSerAsnSerProValSerPro                  275
          :     :::
Db  856 GAGCGCGAGATGAGCGGCGAAGAGCCGGAAGAGCCCTGTACGGCCGCGCGCCC         898
```

RESULT 14
US-09-613-182-5

Qy   67 GlyTyrAspPheAspGlyCysArgLeuArgValGluIle 79
           ||| ||| :::
Db   395 AATGTGTCTTTTGGTTASAACCGGCGGCAGGYTCTCTGTG 433

Search completed: March 26, 2004, 20:26:38
Job time : 117 secs

Qy   178 AlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSer 197
           :::   ::::: ||| ::: |||
Db   716 ------GTCAACCAGAAGAGATTGAGGGCAAGATCGTGCACGTGGCCCCCTTCCAGAAG 769

Qy   198 LysSerTyrArgSerArgSerArg-GlyProSerCysSerTyrSerSerLysSe 217
           ::::: |||:::||::| ||| ||||||||
Db   770 CGCGCTGACCGCCCCCAGGGCAAGGACGTTGTACCAACGTGTTCG----------TC 817

Qy   217 rArgSerValSerProAlaArgSerIleSerProArgSerArg---ProLeuSerArgSe 236
           |||:::   :::   ::: ||| ::
Db   818 AAGAACTTGCCGGCCGACATCGGCGACGACGAGCTGGGCAAGATGGCCACCGAGCACGGC 877

Qy   236 rArgSer------------------------------LeuTyrSe 241
           :|||||
Db   878 GAGATCACCACGGCCGGGTGGTCATGAAGGAGGACGACAAGGGCGGCAGCAAGGGCTTCGGCTTC 937

Qy   241 rSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSe 261
           ||||||| |::| || ::   ::
Db   938 ATCAACTTCAAGGACGCCCGAGT--CGGCCGCCAAGTGCCTGG-----AGTACCTGAAC 988

Qy   261 rLysSerArgSerArgSerArgSerAsnSerProValSerPro 275
           : :::||| :::||| :::|||
Db   989 GAGGCGGAGATGAGCGGCAAGACCCCTGTACGCCGGCCCGCCC 1031

RESULT 15
US-09-621-976-2159
; Sequence 2159, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2159
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 227..472
; NAME/KEY: misc feature
; LOCATION: 18,137
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-2159

Alignment Scores:
Pred. No.:   7.67e-05   Length: 488
Score: 146.50   Matches: 31
Percent Similarity: 57.53%   Conservative: 11
Best Local Similarity: 42.47%   Mismatches: 26
Query Match: 10.10%   Indels: 5
DB: 4   Gaps: 1

US-10-014-927-19 (1-279) x US-09-621-976-2159 (1-488)

Qy   7 ArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeu 26
           ||| :::|| ||||||| ::::::|||::: ||||
Db   230 AGGCGGGTGTTCTGCGGSAACTTCGACGSACGACACCCGCCAGTATGACCTCGAGCGCCTC 289

Qy   27 PheTyrArgTyrGlyProIleValAspIleAspLeuLysIleProArgProProGly 46
           ||| ||||||| |||| ||| ||:::||:::
Db   290 TTCTCCAAGTACGGGCCCCATCAGCCGGATCAGCCATGAAGCTA---------GGC 334

Qy   47 TyrAlaPheValGluPheGluGluProArgAspAlaAspAlaIleTyrGlyArgAsp 66
           |||||::|: |||||||||:||::: ||
Db   335 TATGCTTTCATTTACTTTGAAGATGAGCGTGAGCAAGAGATGYTATAAGGGCTCCGAA 394

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:    March 26, 2004, 19:38:06 ; Search time 432 Seconds
                                       (without alignments)
                                     2404.444 Million cell updates/sec

Title:         US-10-014-927-19
Perfect score: 1451
Sequence:      1 MSSRWNRTIYVGNLPGDIRK.........RSKSRSRSRSNSPVSPVISG 279

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext  0.5
               Ygapop 10.0 , Ygapext  0.5
               Fgapop  6.0 , Fgapext  7.0
               Delop   6.0 , Delext   7.0

Searched:      258946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters:    4917892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10014927/runat_24032004_152927_10016/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10014927 @CGN 1_392 @runat 24032004_152927_10016
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  Published_Applications_NA:*
  1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
  2:  /cgn2_6/ptodata/2/pubppna/PCT_NEW_PUB.seq:*
  3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
  4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
  5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
  6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
  7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
  8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
  9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
 10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
 11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
 12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
 14:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
 15:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
 16:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 17:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 18:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	%	Query Match Length	Pred No.	DB ID	Description
1	1272	87.7	762	9	US-09-938-842A-1199	Sequence 1199, Ap
2	1272	87.7	762	11	US-09-938-842A-1199	Sequence 1199, Ap
3	886	61.1	1209	12	US-10-424-599-134224	Sequence 134224,
4	875	60.3	4044	13	US-10-014-927-18	Sequence 18, Appl
5	854	58.9	1263	12	US-10-425-114-25234	Sequence 25234, A
6	853	58.8	1383	12	US-10-424-599-112845	Sequence 112845,
7	845	58.2	1187	12	US-10-424-599-20355	Sequence 20355, A
8	845	58.2	1320	12	US-10-425-114-24783	Sequence 24783, A
9	824	56.8	1690	12	US-10-425-114-25745	Sequence 25745, A
10	790	54.4	1438	12	US-10-424-599-83457	Sequence 83457, A
11	768.5	53.0	1491	12	US-10-424-599-83461	Sequence 83461, A
12	582.5	40.1	3299	14	US-10-006-285-513	Sequence 513, App
13	542.5	37.4	1717	14	US-10-171-581-9	Sequence 9, Appli
14	517	35.6	1069	9	US-09-880-107-3342	Sequence 3342, Ap
15	517	35.6	1162	9	US-09-925-300-661	Sequence 661, App
16 (c)	517	35.6	1475	15	US-10-264-049-273	Sequence 273, App
17	440.5	30.4	3697	13	US-10-002-600-58	Sequence 58, Appl
18	431	29.7	2167	15	US-10-439-703-38	Sequence 38, Appl
19	418.5	28.8	1534	13	US-10-044-090-268	Sequence 268, App
20	416.5	28.7	2212	9	US-10-925-297-142	Sequence 142, App
21	412.5	28.4	759	9	US-09-910-943-651	Sequence 651, App
22	409	28.2	1579	15	US-10-388-934-197	Sequence 197, App
23	407.5	28.1	446	14	US-10-060-036-52	Sequence 52, Appl
24	388	26.7	345	9	US-09-770-791-783	Sequence 783, App
25	344	23.7	746	9	US-09-910-943-530	Sequence 530, App
26	325	22.4	489	10	US-09-918-995-23144	Sequence 23144, A
27	295.5	20.4	715	9	US-09-910-943-212	Sequence 212, App
28	295.5	20.4	2601	14	US-10-101-510-334	Sequence 334, App
29	279	19.2	742	9	US-09-910-943-539	Sequence 539, App
30	275.5	19.0	282	11	US-09-864-408A-343	Sequence 343, App
31	273	18.8	307	12	US-09-294-093B-1314	Sequence 1314, Ap
32	268.5	18.5	1209	12	US-10-425-114-23865	Sequence 23865, A
33	268.5	18.5	1285	12	US-10-425-114-26648	Sequence 26648, A
34	265	18.3	1252	12	US-10-425-114-24065	Sequence 24065, A
35	265	18.3	1302	12	US-10-425-114-15470	Sequence 15470, A
36	264	18.2	1219	12	US-10-425-114-33401	Sequence 33401, A
37	254	17.5	1322	12	US-10-425-114-19768	Sequence 19768, A
38	244	16.8	1750	9	US-09-925-302-316	Sequence 316, App
39	242	16.7	1264	12	US-10-425-114-22637	Sequence 22637, A
40	242	16.7	1301	12	US-10-425-114-21964	Sequence 21964, A
41	239	16.5	2754	15	US-10-104-047-500	Sequence 500, App
42 (c)	234	16.1	3026	10	US-09-919-039-314	Sequence 314, App
43	231	15.9	432	9	US-09-919-039-314	Sequence 1284, Ap
44	231	15.9	432	13	US-10-033-528-1284	Sequence 1284, Ap
45	231	15.9	432	14	US-10-099-926-1284	Sequence 1284, Ap

ALIGNMENTS

```
RESULT 1
US-09-938-842A-1199
; Sequence 1199, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1199
; LENGTH: 762
; TYPE: DNA
```

```
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1199
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1199
```

```
Alignment Scores:
Pred. No.:              9.88e-123      Length:         762
Score:                  1272.00        Matches:        253
Percent Similarity:     91.67%         Conservative:   0
Best Local Similarity:  91.67%         Mismatches:     1
Query Match:            87.66%         Indels:         23
DB:                                    Gaps:           2
```

```
US-10-014-927-19 (1-279)  x  US-09-938-842A-1199 (1-762)

Qy    1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyValAspIleArgLys  20
Db    1 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG     60

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle     40
Db   61 TGTGAGGTTGAAGATCTCTTCACAAGTATGTGGACCAATTGTGACATTGATTTGAAGATT    120

Qy   41 ProProArgProProGlyTyrAlaPheValGluValAspProArgAlaAspAlaAspAsp     60
Db  121 CCACCCGAGACCTCCTGGTTATGCCTTTGTCGAGTTTGAAGATCCTCGTGATGCAGACGAT   180

Qy   61 AlaIleTyrGlyArgArgAspGlyTyrArgAspPheSerAspGlyCysArgLeuArgValGluIleAla  80
Db  181 GCAATTTATGGACGTAGGGATGGTTATGATTTTGATGGGTGTCGACTTCGGGTTGAGATTGCA 240

Qy   81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer      100
Db  241 CATGGTGGTCGTAGATTTCACCATCAGTTGATAGGTACAGCAGCAGCTACAGTGCCAGC    300

Qy  101 ArgAlaProSerArgArgAspTyrArgValLeuValThrGlyLeuProProSerAla      120
Db  301 CGTGCACCTTCAAGACGCTGTAGATTTTGTCACCGGATTACCGCCTTCTGCT          360

Qy  121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyValCysPheSerGluVal      140
Db  361 TCGTGGCAGGACCTTAAGGATCACATGCGCAAAGCTGGAGTGTCTGCTTCTGAAGTT      420

Qy  141 PheProAspAlaArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db  421 TTCCCTGACGCTAGAAAGGGCATGTCTGGGGTTGTGGATTATAGCAACAATGATGATGAAG  480

Qy  161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerArgAlaTyrIle 180
Db  481 TACGCAATAAGGAAACTTGATGCCACTGAATTTCGAAATGCTTTCTCTAGTGCTTATATA   540

Qy  181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspSerProValSerSerTyr 200
Db  541 CGGGTGAGGGAATATGAGTCGAGGAGTGTGAGTCGAAGCCCAGATAGTCCTGTA        600

Qy  201 ArgSerArgSerArgSerArgGlyArgSerArgGlyProSerCysSerSerTyrSerArgSerVal 220
Db  601 AGAAGCAGGAGTCGG------------------------------AGCCGTGT-GTG      626

Qy  221 SerProAlaArgSerIleSerProArgLeuSerProArgSerArgSerArgLeuTyr       240
Db  627 TCACCTGCTAGATCCAGATCCATTTCCCCGGTTCACGGCCCCTAGTCGTTCGTCGCTATAC  686

Qy  241 SerSerValSerArgSerGlySerLeuLeuArgAlaGlySerLeuAspTrpIleLeuSerGlnSerArg 260
Db  687 AGCTCTGTCTCAAGG------------------------------TCCCAATCAAGA      713

Qy  261 SerLysSerArgSerArgSerArgSerArgSerAsnSerProValSerProVal         276
Db  714 TCAAAATCAAGATCAAGATCAAGATCTCCAGTTCTCCAGTTTCACCTGTA            761
```

```
RESULT 2
US-09-938-842A-1199
; Sequence 1199, Application US/09938842A
; Publication No. US2004000476A9
```

```
Qy   201 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220        Qy    99 AlaSerArgAlaProSerArgAspTyrArgValLeuValThrGlyLeuProPro 118
             ||||||||||                                     |||||             373 GGTAGCCGTGGAGTTCCAGGCCGTTCTCGACTACCGTGTTCTCGTCACTGGATTACCTCCG 432
Db   601 AGAAGCAGGAGTCGG--------------------------AGCCGTGT-GTG 626        Db

Qy   221 SerProAlaArgSerIleSerProSerArgProLeuSerArgSerArgSerLeuTyr 240      Qy   119 SerAlaSerTrpGlnAspLeuLysAsnHisMetArgLysAlaGlyAspValCysPheSer 138
             |||||||||||||                                                       |||||||||||||
Db   627 TCACCTGCTAGATCCATTCCCCGCGTTCACGGCCCCTTAGTCGTTCTCGCTCGCTATAC 686    Db   433 TCTGCTTCATGGCAAGACCTGAAAGATCACATGCGTAAAGCTGGTGATGTGTGTTTTTCG 492

Qy   241 SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArg 260    Qy   139 GluValPheProAlaArgLysGlyMetSerGlyValValAspTyrSerAsnTyrArgAspAsp 158
             |||||||||                                                            ::: |||||||:::    :::     ||||:::
Db   687 AGCTCGTGTCTCCAAGG----------------------TCCCAATCAAGA 713            Db   493 CAAGTCTTCCGTGAGCGTGGGGTATGCGTGGGATAGTGATAGTTGATTATACCAATTATGATGAT 552

Qy   261 SerLysSerArgSerArgSerArgSerAsnSerProValSerProVal 276               Qy   159 MetLysTyrAlaIleArgGlyLeuAspAlaThrGluAspArgAsnAlaPheSerSerAla 178
             ||||||||                                                             |||                :::||||
Db   714 TCAAAATCAAGATCAAGATCGAATTCTCCAGTTTCACCTGTA 761                    Db   553 ATGAAATATGCTATCAGGAAACTTGATGACTCAGAATTCCGCAATGCCTTTCTCGTGCT 612

                                                                                  Qy   179 TyrIleArgValArgGluTyrGluSerArgValSerArgSerPro------AspAsp 196
RESULT 3                                                                                  |||         |||    :::|||
US-10-424-599-134224                                                        Db   613 TTCATACGGGTGAGGGAATATGAT---CGAGGTTATTCTAGAAGTCCTAGTCGTGATTCA 669
; Sequence 134224, Application US/10424599
; Publication No. US20040031072A1                                          Qy   197 SerLysSerTyr                                   ---ArgSerArgSerArg 205
; GENERAL INFORMATION:                                                                :::||||||
; APPLICANT: La Rosa Thomas J                                              Db   670 AGGAGGAGCTATTCTAGAAGCCTCAGCCGCAGTCCCATGTATCTCACGAAGCCGCAGCCAT 729
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua                                                    Qy   206 SerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSer 225
; APPLICANT: Cao Yongwei                                                               ||||||            :::|||||
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules       Db   730 AGCCGCAGCAGAGTCATAGCTATAGTGACAGGAGCAGAGTTGGTCTCCGAAACCAAAA 789
;   Associated With Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B                                            Qy   226 IleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSer 241
; CURRENT APPLICATION NUMBER: US/10/424,599                                           |||||             :::||||
; CURRENT FILING DATE: 2003-04-28                                         Db   790 CACTCCTGCAGCGTTCAATATCTCTTTCAAGGTCAGTTTCTACCAAGTCT 837
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 134224
; LENGTH: 1209                                                           RESULT 4
; TYPE: DNA                                                              US-10-014-927-18
; ORGANISM: Glycine max                                                  ; Sequence 18, Application US/10014927
; FEATURE:                                                               ; Publication No. US2002011518OA1
; NAME/KEY: unsure                                                       ; GENERAL INFORMATION:
; LOCATION: (1)..(1209)                                                  ; APPLICANT: Barta, Andrea
; OTHER INFORMATION: unsure at all n locations                           ; APPLICANT: Lopato, Sergyi
; FEATURE:                                                               ; APPLICANT: Kalyna, Maria
; OTHER INFORMATION: Clone ID: PAT_MRT3847_92209C.1                       ; APPLICANT: Dorner, Silke
US-10-424-599-134224                                                     ; TITLE OF INVENTION: Splice Factor
                                                                         ; FILE REFERENCE: SONN:013US
Alignment Scores:                                                        ; CURRENT APPLICATION NUMBER: US/10/014,927
Pred. No.:                                                               ; CURRENT FILING DATE: 2001-10-23
Score:               886.00                                              ; PRIOR APPLICATION NUMBER: PCT/AT00/00100
Percent Similarity:  78.52%                                              ; PRIOR FILING DATE: 2000-04-20
Best Local Similarity: 72.27%                                            ; PRIOR APPLICATION NUMBER: A 727/99
Query Match:         61.06%                                              ; PRIOR FILING DATE: 1999-04-23
DB:                  12                                                  ; NUMBER OF SEQ ID NOS: 22
            2.15e-82   Length:    1209                                   ; SOFTWARE: PatentIn Ver. 2.1
                       Matches:   185                                    ; SEQ ID NO 18
                       Conservative: 16                                  ; LENGTH: 4044
                       Mismatches: 39                                    ; TYPE: DNA
                       Indels:    16                                     ; ORGANISM: Unknown
                       Gaps:      4                                      ; FEATURE:
                                                                         ; OTHER INFORMATION: Description of the unknown organism:genome
US-10-424-599-134224 (1-279) x US-10-424-599-134224 (1-1209)            ; OTHER INFORMATION: atSRp30
                                                                        US-10-014-927-18
Qy     1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
            |||||||||||                         :::|||              :::||       Alignment Scores:
Db    73 ATGAGCAGCGGGCGGTCAAGACCCGCACAATTTATGTTGGCAATCTTCCTGGCGACGTTCCGCTTG 132  Pred. No.:
                                                                         Score:               875.00           Length:      4044
Qy    21 CysGluValGluAspPheLysTyrLysTyrGlyTyrProIleValAspIleAspLeuLysIle 40     Percent Similarity:  43.53%           Matches:     242
            ||||||||||                                                            Best Local Similarity: 43.53%         Conservative: 0
Db   133 AGAGAAGTAGAGGAATCTATTCTCACAAGTATGCTCCATTGTTGACATTGATTTGAAGATA 192        Query Match:        60.30%            Mismatches:  3
                                                                         DB:                  13               Indels:      314
Qy    41 ProProArgProGlyTyrAlaPheValGluGluGluProArgAspAlaAspAsp 60                            1.41e-80             Gaps:        9
            |||||||                                    :::||
Db   193 CCTCCGAGACCACCAGGTATGCTTTTGTTGAGACTTTGAGGATGATGCTCGTGATGCTGAAGAT 252     US-10-014-927-19 (1-279) x US-10-014-927-18 (1-4044)

Qy    61 AlaIleTyrGlyArgAspGlyTyrArgAspPheAsnGlyCysArgLeuArgValGluIleAla 80      Qy     1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
            ||||||                         :::|||                                      ||||||||||||                          :::||           :::||
Db   253 GCAATTCAATATCGAGATGGTTACAATTTGATGGTTTCGGTTACGAGTTGAACTTGCA 312         Db   808 ATGAGTAGCCGATGGAATCGAACTACCTTGGGAATTTGCCTGGCGGAATATTCGCAAG 867

Qy    81 HisGlyArgGlyArgAspPheSerProSerValAspArgTyrSer-----SerSerTyrSer 98
            ||||||                         :::||
Db   313 CATCGTGGACGCGGGGATATTCATCATCAGTAGAACCGTTACAGTAGCTATAGCGGTGGCAGT 372
```

```
Qy   21 CysGluValGluAspLeuPheTyrLys--------------------- 29
Db  868 TGTGAGGTTGAAGATCTCTTCACAAGGTTTGAAAATTTCCTCTTTTCTCGATAAAA 927

Qy   29 ----------------------------------------------- 29
Db  928 TTGAATTCATTATGACTAGTTGGGTTCATAAATTCGTCTTCTGCTGAGACAAT 987

Qy   30 -------------TyrGlyProIleValAspIleAspLeuLy------- 39
Db  988 TTAAAATCGACTCTTATGTATATTTGTTTCAGTATGACGACCAATTGGGACAATTGATTTGAA 1047

Qy   39 sIleProProArgProGlyTyrAlaPheValGlu-------------- 51
Db 1048 GATTCCACCGAGACCCTCGGTTATGCCTTTGTCGA-GGTATATTGATCAAGTACAAATT 1106

Qy   51 ------------------------------------------------ 51
Db 1107 TGTTTTTTTCTTCTTGTAATATAGGCTATAATGACTAAGATAGTTGTTATTGGTG 1166

Qy   52 ----PheGluAspProArgAspAlaAspAlaIleTyrGlyValArgAspGlyTyrAspPh 70
Db 1167 GCAGTTTGAAGATCCCGTGATGCAGACGAACGTGATGGTTATGATTT 1226

Qy   70 eAspGlyCysArgLeuArg------------------------------ 76
Db 1227 TGATGGGGTGTCGACTTCGGGTTAGTAAACGCATGAGAAAGCTAGCTTAATTTCTGTAA 1286

Qy   77 ------------------------ValGluIleAlaHisGlyGlyAla 84
Db 1287 TTTCTTGTAAAGGTGTTATCTTTGTGTGATGTTTTTTAGGTTGAGATTGCACATGGTGGTC 1346

Qy   84 rgArgPheSerProSerValAspArgTyrSerSerAlaSerArgAlaProS 104
Db 1347 GTAGATTTTCACCATCAGTTGATAGGTACAGCAGCAGTACGACGCCGAGCCGTGCACCTT 1406

Qy  104 erArgSerAspTyrArg------------------------------- 110
Db 1407 CAAGACGCCTCGCACTACCGCGCCGGTTGTAGAGTCTTCTCGATTGTGTTATTTGGTGTTGTG 1466

Qy  111 -----------------------------------ValLeuValThrGly 115
Db 1467 TAAAATTTTATATTTGAAAACTCATTTTACTACCTAAACATGTAGTGCTTGTGACCGGA 1526

Qy  116 LeuProProSerAlaSerTrpGlnAspLeuLys---------------- 126
Db 1527 TTACCGGCCTTCGTGCTTCGTGCGTGGCAGGACCTAA-GGTAAGGGACACTATATAGTCTTTTC 1585

Qy  126 ------------------------------------------------ 126
Db 1586 TCTGAATGTGGTTCCTATATCATGTTTTGGATTTATCTCTTTTCTGAAATGATGTTAT 1645

Qy  127 ----------------AspHisMetArgLysAlaGlyIleAspValCysPheSerGlu 139
Db 1646 TTGCTATTACGGGTGATTAGGATCCACCATCGCCAAAGCTGGAGATGTCTGCTTCTTCTGAA 1705

Qy  140 ValPheProAspArg-Lys-------------- 145
Db 1706 GTTTTCCCTGACCGTAAAGGTGAGTGAGTTAGCGATATTCGGATAGTTTGGATAAGCTTTTTTGATTGA 1765

Qy  145 ------------------------------------------------ 145
Db 1766 TGTGTTAGTAAATTAGTCTTTGTGAAGGAGAATAGTGTTAAGCATCTGAACTGCTAAAC 1825

Qy  146 --------GlyMetSerGlyValValValAspTyrSerAsnTyrArgAspAs 158
Db 1826 TCACATTCAGTATTCTTTGTAGGCATGTCTGGGGTTGTTGTGATTATAGCAACTATGATGA 1885

Qy  158 pMetLysTyrAla-------------------- 162
Db 1886 TATGAAGTACGC-AGTAAGTTTATATCTTTGCAACGCAAATGTTCCTGGACTTATGCCT 1944

Qy  162 ------------------------------------------------ 162
Db 1945 TAGACTGCTTTGTTCATAGTACACGGAGCTGAATTTATCTTCCTGGAGGCCAGTGTTG 2004

Qy  163 --------------------------IleArgLysLeuAspAlaThrGl 170
Db 2005 GATCTTGATGTTCCCTAAAATTTTGATGTGACAGATAAGGAAACTGATGCCACTGA 2064

Qy  170 uPheArgAsnAlaPheSerAlaTyrIleArg----------------- 181
Db 2065 ATTTCGAAATGCTTTCTCCAGTGCTTATATACGGGTATGTTGTATTGCTTCTTCTTGATTT 2124

Qy  182 --------------------------------ValArgGluTyrG 186
Db 2125 TGTTAAGCATAAGTGGATATGGAGTCATCTCTGAATTACTGTTCAGTGAGGGAAATATG 2184

Qy  186 1uSerArgValSerArgSerProAspAspSerLysSerArgSerArgs 206
Db 2185 AGTCGGAGGAGTGTGAGTCGGAAGCCCCAGATGATTCTAAAAGCTATAGAAGCAGGAGTCCGGA 2244

Qy  206 erArgGlyProSerCysSerTyrSerSerSerLysSerArg-------- 218
Db 2245 GCCGTGGTCCAAGCTGTAGCTAAGTAGCAAGAGCAGAGGGTCTTTTTTTTTTTTTTT 2304

Qy  218 ------------------------------------------------ 218
Db 2305 TTCATAAACCTAAGACATATAAGGAGATTTTATTGTAACTATTATTATGAAAATTAACTGAC 2364

Qy  219 ------------SerValSerProAlaArgSerIleSerProSerArgProLeu 233
Db 2365 TTCTAAATGCAATGCAGTGTGTCACCTGCTAGATCCATTCCCCGCGTTCACCGGCCCCTT 2424

Qy  234 SerArgSerArgSerLeuTyrSerSerValSerArg 245
Db 2425 AGTCGTTCCCGCTGCTATACCAGTCTCGTCTCAAGG 2460
```

RESULT 5

```
Alignment Scores:
Pred. No.:                                    Length:         1263
Score:                  4.78e-79              Matches:         183
Percent Similarity:     854.00                Conservative:     23
Best Local Similarity:  76.30%                Mismatches:       48
Query Match:            67.78%                Indels:           16
DB:                     58.86%                Gaps:              4
                        12
```

US-10-425-114-25234 (1-279) x US-10-425-114-25234 (1-1263)

```
Qy 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgGlyLys 20
```

```
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_72906C.1
US-10-424-599-112845

Alignment Scores:
Pred. No.:                                    1383
Score:             6.82e-79
                   853.00        Matches:        190
Percent Similarity:       69.38%   Conservative:   23
Best Local Similarity:    61.89%   Mismatches:     58
Query Match:              58.79%   Indels:         36
DB:                                Gaps:           6

US-10-014-927-19 (1-279) x US-10-424-599-112845 (1-1383)
```

```
Qy    1  MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyValAspIleArgLys   20
         |||||                     |::|||||||||||||||||||||||||||||:::
Db  153  ATGAGTAGCAGCCGCTCGAGCCAGAGAACTGTCTCACGTTGGGAATCTACCCGGTGATATCCGTGAA  212

Qy   21  CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle   40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||:::||
Db  213  AGAGAAGTTGAAGATTGTTTCTGAAGTACGGACCACATAACACCACATCGACCTAAAGGTT  272

Qy   41  ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgArgAlaAspAsp   60
Db  273  CCCCCAAGACCACCTGGTTACGCATTTGTAGAGTTTGAAGATGCTCAAGATGCTGAGGAT  332

Qy   61  AlaIleTyrGlyArgArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla   80
Db  333  GCAATTCGTGGTCGTGATGGCCATGATTTTGATGGCCACCGGTTACGGGTGGAGCCTGCC  392

Qy   81  HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSer------   98
         :::
Db  393  CATGTGGGGCGTGGTCATCATCCTCAAAAGATGTCCAAAATAGTCACGCAATGGCCGA  452

Qy   99  AlaSerArgAlaProSerArgSerAspTyrArgValLeuValThrGlyValLeuProPro  118
Db  453  GGTGACGTGGGGTATCTAGGCGCTCGAATATCGTGTTCTAGTCACTGGATTGCTTTCT  512

Qy  119  SerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaAlaGlyAspValCysPheSer  138
Db  513  TCGGCCATCCTGGCAGGATCTTAAGGATCACATGCGGCAGAGCCAGGGGATGTTTGCTTTTCT  572

Qy  139  GluValPheProAlaArgLysGlyMetSerGlyValValAlaSerAsnTyrArgAsp  158
Db  573  CAAGTTTTTCATGATGGAAGGGGTACTACCGGTATTGTGATTACCAAATTATGATGAT  632

Qy  159  MetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAla  178
         :::||||||                  |||
Db  633  ATGAAATATGCTATTAAGAAGCTTGATGACTCGAGTTCCGGAATGCATTTTCCAAGGGA  692

Qy  179  TyrIleArgValArgGluTyrGluSer---ArgSerValSerArgSerProAspAspSer  197
Db  693  TATGTTCGGGTGAGGGAATATGATTCAAGGCGGGATTCTTCTAGAAGTCCTAGTCATGGC  752

Qy  198  LysSerTyr--------ArgSerArgSerArgGlyProSerCysSerTyrSer  214
Db  753  CCATCTCATTCTAGAGGATGGAGGAGCTATAGCCGACCGTAGCCGCATAGTTACAGT  812

Qy  214  ----------------------------------------  214
Db  813  CGGGACCGGAGCCAAGCAAATCTCCAAAAGGAAAGTCTTCACAGCGTTCACCTGCTAAA  872

Qy  215  --------SerLysSerArgSerValSerProAlaArgSerIleSer------Pro  228
Db  873  TCTCCAAAAGGAAAGTCTTCACAGCGTTCACCTGCTAAATCCATCAAGATCTCCCTCT  932

Qy  229  ArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgGlySer  248
Db  933  CGTTCTCGCTCAAGTCAAGGTCCCGTTCGTTATCAGGATCACGGTCCAGGTCT  992

Qy  249  LeuLeu----------ArgAlaGlyAspTrpIleSerGlnSerArgSerArg  264
         :::::
Db  993  CCATTGCCTCCGCGTAATAAAAAGCCCCAAAAAACGCAGTGCTAGCAAGTAGG  1052
```

```
Db   72  ATGAGCAGGCGCAACAGCCGCTACCATCTATGTAGGCAATCTCCCTGGGACATCCGTGAG  131
Qy   21  CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle   40
Db  132  AGGGAGGAGGTTGAGGGATCTCTTCACAAGTATGGCCGTATTTTGGATATTGACTTGAAAATA  191
Qy   41  ProProArgProProGlyTyrArgAlaPheValGluPheGluAspProArgArgAlaAspAsp   60
Db  192  CCTCCGAGACCTCCTGGATACGCATTCGTTCGTTGAGTTTGAGGAGGATCCACGTGATGCTGATGAT  251
Qy   61  AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla   80
Db  252  GCAATTTATGCCGTGATGGGTATAACTTTGATGGCTACAGGTTGAGGGTTGAATTAGCT  311
Qy   81  HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer  100
Db  312  CATGGTGGCGCAGAGGCCAGTCTTATTCTTATCGTTCTATGATCGTTCAAGCAGTGCATGC  371
Qy  101  Arg---AlaProSerArgSerAspTyrArgValLeuValThrGlyValLeuProProSer  119
Db  372  CGTGGAGGTGTTTCTAGGCGCCTCTGATTTCCGTGTTATGGTTCACTGGTTACCCTCATCG  431
Qy  120  AlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaAlaGlyAspValCysPheSerGlu  139
Db  432  GCATCGTGGCAAGATCTGAAGGACCACATGCGGCAGAGCCAGGGGATGTTTTGCTTTTCTGA  491
Qy  140  ValPheProArgLysGlyMetSerGlyValValAspTyrArgSerAsnTyrArgAspMet  159
Db  492  GTATACCGTGAGGTCTGAGAACTATGGAATTGTGATTATACAAATTATGACGATATG  551
Qy  160  LysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyr  179
Db  552  AAATACGCGATTAGGAAGCTTGATGACTCACAGTTCAAGAATGCATTTCAAGAGCATAT  611
Qy  180  IleArgValArgGluTyrGluSerArgSerArgSerValSerArgSerProAspSerLysSer  199
Db  612  ATCAGGGTGAGGGAGTATGAGTCTAGTAGATCACGAAGCAGAAGC-----------  653
Qy  200  TyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSer  219
Db  654  ---CGTAGCCACTCGTACTCTAGAAGCCCCAGCTACAGGAGCAGGAGTCCAAAATCT  710
Qy  220  ValSerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeu  239
Db  711  GTTTCTCAGTCACCCTCA---TCTGTGGATGAAAGATCGCTATACAAGATCTCGATCCCCA  767
Qy  240  TyrSerSerValSerArgGlySerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSer  259
Db  768  ATTTCTTCTCCTTCTCATGCA---------AGATATGGCAACAAGCCCT  806
Qy  260  ArgSerLysSerArgSerArgSerArgSer  269
Db  807  AGGAGCAGGAAGAGCGCATCCCGGTTCTCGGTCT  836
```

```
RESULT 6
US-10-424-599-112845
;  Sequence 112845, Application US/10424599
;  Publication No. US20040031072A1
GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT: Kovalic David K
;  APPLICANT: Zhou Yihua
;  APPLICANT: Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE: 2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 112845
;  LENGTH: 1383
;  TYPE: DNA
;  ORGANISM: Glycine max
```

Qy   265 SerArgSerArgSerAsnSer 271
              |||||||||||||
Db   1053 AGCAGGAGCAGGAGCAAGAGT 1073

RESULT 7
US-10-425-114-20355
; Sequence 20355, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20355
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3180-041-E2_FLI
US-10-425-114-20355

Alignment Scores:
Pred. No.:   3.79e-78   Length: 1187
Score: 845.00   Matches: 186
Percent Similarity: 73.05%   Conservative: 20
Best Local Similarity: 65.96%   Mismatches: 44
Query Match: 58.24%   Indels: 32
DB: 12   Gaps: 8

US-10-014-927-19 (1-279) x US-10-425-114-20355 (1-1187)

Qy   1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
             |||:||:||||| || |||| ||| :::
Db   73 ATGAGCAGGCGCTGGAGCGCGCACGATCTACGTCGGGAACCTCCCCGGCGACATCCGGGAG 132

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIle 40
             ||| | || | |||| :::
Db   133 AGGGAGGTGGAGGATCTGTTCTACACAAGTAAGATGGTAAAATTGTTGACATTGACCTGAAGGTC 192

Qy   41 ProProArgProProGlyTyrAlaPheValGluAspGluAspProArgAspAlaAspAsp 60
             | ||| ||||| |||
Db   193 CCCCCAAGACCACCTGGTTATGCTTTGTTTGTTGAGTTGAAGATCCTCGTGATGCTGAGGAG 252

Qy   61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
             |||| || ||| ||
Db   253 GCAATTGCTGGACACGGGATGGATGAACTTGATGGACACCGTCTAAGAGTGGAGGCTGCT 312

Qy   81 HisGlyGlyArgArgPheSerProSerValAspArg----------TyrSerSerSerTyr 97
             ||||||| | ::: ||| || ||| :::
Db   313 CATGTGTGGTAGAGGTAAATGCTTCCTCCGCATGATCGTTCAAGTGGCTTTGGTGCGGTGGT 372

Qy   98 SerAlaSerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuPro 117
             ||||| | ||||
Db   373 GGAGCACGTCGTGGTGGTCGAGAGCACTCAGAGTATCGTGTTCTTGTCACTGGACTGCCT 432

Qy   118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysMetAlaGlyValCysPhe 137
             |||| | ||| |||||
Db   433 TCTTCTGCATCATGGCAGGATTTAAAGGATCATATGCCGAAGGCTGGTGATGTTTGTTTC 492

Qy   138 SerGluValPheProAspArgLysMetSerGlyValValAspTyrSerAsnTyrAsp 157
             ||||||:| | ::: || ::: ||||||
Db   493 TCTGAAGTGTATCGCCGAAGGCGGTCGGCACCGTGGACTACCACAAATTATGAT 552

Qy   158 AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer 177

Alignment Scores:
Pred. No.:   9.06e-76   Length: 1690
Score: 824.00   Matches: 187
Percent Similarity: 58.50%   Conservative: 23
Best Local Similarity: 52.09%   Mismatches: 45
Query Match: 56.79%   Indels: 104
                                                 DB: 12   Gaps: 5

US-10-014-927-19 (1-279) x US-10-425-114-25745 (1-1690)

```
Qy      1  MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys  20
              :::||  ||||||||||||||||||||||||||||||
Db    145  ATGAGCAGGCGCAACAGCGCTACCATCATGATGAGGCAATCTCCCTGGGACATCCGTGAG  204

Qy     21  CysGluValGluAspPheTyrLysTyrGlyTyrProIleValAspIleAspLeuLysIle  40
              |||||||||||||||||||||||||||  :::||||||||||||||||
Db    205  AGGGAGGTTGAGGATCTTCTTCTACCAAGATGGCCGTATTTGGATATTGACTTGAAAATA  264

Qy     41  ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp  60
              ||||||||||||||| :: :|||||||||||||||||||||||||
Db    265  CCTCCGAGACCTCCTGGATACGCATTCGTTGAGTTTGAGGATCCACGTGATGCTGATGAT  324

Qy     61  AlaIleTyrGlyTyrAspPheAspGlyCysArgLeuGlyArgValGluIleAla  80
              ||||||| |||||||||||||||| :::|||||||||||||||||
Db    325  GCAATTTATGGCGTGACGGGTATAACTTTGATGGCTACACAGGTTGAGGGTTGAATTAGCT  384

Qy     81  HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerAlaSer  100
              :::  |||||| :::|||
Db    385  CATGGTGGCAGAGGCCAGTCTATTCTTATCTTATGATCGTTCAAGCAGTAGCAGTGCATGC  444

Qy    101  Arg---AlaProSerArgSerArgProArgSerArgValLeuValThrGlyLeuProProSer  119
             ||| |||||||||||||||||||||||||
Db    445  CGTGGAGGTGTTTCTAGGGGCTCTGATTTCCGTTTATGTCACTGGTTTACCCTCATCG  504

Qy    120  AlaSerTrpGlnAspLeuLysAspHisMetSerGlyValCysPheGlu  139
             |||||||||||||||||| | |||||||||||||||||
Db    505  GCATCGTGGCAAGATCTGAAGGACCCACATGCGCGCGCTGGTGATGTCTGTTTCTCGAT  564

Qy    140  ValPheProAspArgLysGlyMetSerGlyValValAspTyrArgAspAspMet  159
              ||: ||:||||||||||||||| ::|||
Db    565  GTATACCGTGAGGCTGAGGAAACTATTGGAGTTGGATTATACAAATTATGACGATATG  624

Qy    160  LysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyr  179
              |||||||||||| :::|||||||
Db    625  AAAATACGCGATTAGGAAGCTTGATGACTCCAGTTCAGGAATGCATTTTCAAGAGCATAT  684

Qy    180  IleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSer  199
              ||| :::||||
Db    685  ATCAGGGTGAGGGAGTATGATGCTAGATCACGAGCAGAAGC----------  726

Qy    200  TyrArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSer  219
              ||||| :::
Db    727  ---CGTAGCCACTCGTACTCTAGAAGCCCAGTCTACAGCAGGAGCAGGAGTCCAAAATCT  783

Qy    220  ValSerProAlaArgSer  225
             |||||| :::
Db    784  GTTTCTCAGTCACCTCATCGTGGATGAAAGGTTGCTTTGTTTTTCCTAGATTTTGGC  843

Qy    225  -------------------  225
Db    844  CATGCGGCGATCTGGGGCTGCATTGGAATCGGAGTGGA  903

Qy    225  -------------------  225
Db    904  TGGTTTGGGACTTCGGATAGTAGAGATGATTGTGATGCATGTATGATGGGCTACA  963

Qy    226  IleSerProArgSer  230
             |||||||
Db    964  ATCAGTCCTAGGAGCAAGATAGGCTGACACGTAAGAAAATCTGAAGTAAGCTTGTGGTGT  1023

Qy    230  -------------------  230
```

```
Qy     61  AlaIleTyrGlyTyrAspPheAspGlyCysArgLeuGlyArgValGluIleAla  80
              |||||||
Db    364  GCAATTCTGGACGCGGATGGATACAACTTTGATGGACACCGTCTAAGAGTGGAGGCTGCT  423

Qy     81  HisGlyGlyArgArgPheSerProSerValAspArg-------TyrSerSerSerTyr  97
              ||||||| :::
Db    424  CATGGTGGTAGAGGTAATGCTTCCTCGCATGATCGTTCAAGTGGCTTTGGTGGCGGTGGT  483

Qy     98  SerAlaSerArgAlaProSerArgSerArgProArgSerArgProArgSerArgValThrGlyLeuPro  117
              |||||| :: :|||||||
Db    484  GGAGCACGTCGTGGTGGTCGAGACACTCAGAGTATCGTGTTCTTGTCACTGGACTGCCT  543

Qy    118  ProSerAlaSerTrpGlnAspLeuLysAspHisMetSerGlyValCysAspValCysPhe  137
              ||||||||||||
Db    544  TCTTCTCCATCATGGCAGGATTAAAGGATCATATGCGGAAGGCTGGTGATGTTTGTTTC  603

Qy    138  SerGluValPheProAspArgLysGlyMetSerGlyValValValAspTyrSerAsnTyrAsp  157
              |||||| :::  |||
Db    604  TCTGAAGTGTTCCGGAAGGCGGTGGCACCGTAGGAATTGTGGACTACACAAATATGAT  663

Qy    158  AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer  177
              ::: |||
Db    664  GATATGAAAATATGCTATAAAGAGCTGGATGATACTGAATTCAGGAACGCCTTTGGGCGA  723

Qy    178  AlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSer  197
             :::
Db    724  GCCTATATAAGGGTGAAGAATAT-------------------AACGGC  753

Qy    198  LysSerTyrArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSer  217
              |||||||
Db    754  AAACGTGGTCGCTCCTACTCACGTAGCCGAAGCCCAAGTCGTAGTTAC---AGCAAAAGC  810

Qy    218  ArgSerValSerProAlaArgSerIleSerProArgSerArg-------ProLeuSer  234
             |||||||
Db    811  AGGAGTCCGAGTAAA-------TCACCCAGGACTCGCCGTTCATCATCATCTAGATCC  858

Qy    235  ArgSerArgSerLeuTyrSerSerArgValSerArgSerGlySerLeuLeuArgAlaGlyAsp  254
             :::
Db    859  CGGTCAAGATCTGTT---TCTTCTCGTTCAAGGTCCCA  894

Qy    255  TrpIleSerGlnSerArgSerArgSerArgSerArgSerAsnSerPro---Val  273
             |||||| :::
Db    895  -------TCAAAAGGACGTTCTCCATCAAGATCCCAAGATCGAAATCTCCTAATGTT  948

Qy    274  SerPro  275
             ||||||
Db    949  TCTCCA  954
```

RESULT 9
US-10-425-114-25745
; Sequence 25745, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25745
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3959-003-F10_FLI
US-10-425-114-25745

Db   617 GACGTGTGTTTGCTGAGGTTCCGTGATAGTGAAGGGACTTTTGGCATTGTTGATTAC 676

Qy   154 SerAsnTyrAspAspMetTyrTyrAlaIleArgGlyLeuAspAlaThrGlnPheArgAsn 173
                 :::::: |||| | || | || ||| |||| :::
Db   677 ACTAATTATGATGACATGAAGTATGCCATTCGGAAATTGGATGACACTGAGTTTAGAAAT 736

Qy   174 AlaPheSerAlaGlyAlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSer 193
                 :::::| | | || |: | | |::|:|||
Db   737 CCCTGGGCGGAGAGCTTATATTCGGTGAGGAAGCATGAGAGTAGT----------------- 781

Qy   194 ProAspAspSerLysSerArgTyrArgSerArgSerArgSerArgGlyProSerCysSerTyr 213
Db   782 -----------------CGTTCTAGGAGCCCCAGCAGGAGGAGCCCC---------------- 808

Qy   214 SerSerLysSerArgSerValSerProAlaArgSerIleSerProArgSerArgProLeu 233
                 ||||| || ||| | | |||| || |:
Db   809 -------------AGAGGAGCCCCAGCAGGACAGAAGTCCAAAAGA-------------- 844

Qy   234 SerArgSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuLeuArgAlaGly 253
                 :::::| |||||||||||||||:||:: |||||||| :::::
Db   845 GTTAGAAGTAGATCCTCGGAGCGTTCTGTGTCGTTAGGTCAGTGTCAATCAGCA 904

Qy   254 AspTrpIleSerGlnSerArgLysSerArgSerArgSerAsnSer---Pro 272
Db   905 TCACCAATCAAATCTTCTAGGCCAAGATCACGATCCGATCACGATCTCGATCT 964

Qy   273 ValSerPro-----ValIleSerGly 279
Db   965 GGATCTCCTCATCAGGTCCTGTCAGGT 991

RESULT 11
US-10-424-599-83461
; Sequence 83461, Application US/10424599
; Publication No. US2004031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 83461
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_46381C.1
US-10-424-599-83461

Alignment Scores:
Pred. No.:
Score: 4.46e-70   Length: 1491
Percent Similarity: 768.50   Matches: 171
Best Local Similarity: 67.12%   Conservative: 25
Query Match: 58.56%   Mismatches: 51
DB: 52.96%   Indels: 45
                            12   Gaps: 7

US-10-014-927-19 (1-279) x US-10-424-599-83461 (1-1491)

Qy   1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
               ||||||| |: || ||| || || |::
Db   418 ATGAGCGGTCGCTTTTCTCGCACAATTTATGTCGGCAACCTGCCCCTGATATAAGAGAG 477

Qy   21 CysGluValGluAspPheTyrLysPheGlyProIleValAspIleAspLeuLysIle 40
               ||| | | ::||||| | ||| || ||
Db   478 TCAGAAATTGAAGATCTATTCTTACAAGTATGCCGTATAATTGAATTGAAGGTT 537

Qy   41 ProProArgProProGlyTyrAlaPheValGluProGlyAspArgAspAlaAspAsp 60
               ||||::|||

Pred. No.:   2.55e-50   Length: 3299
Score: 582.50   Matches: 134
Percent Similarity: 64.80%   Conservative: 28
Best Local Similarity: 53.60%   Mismatches: 47
Query Match: 40.14%   Indels: 41
DB: 14   Gaps: 8

US-10-014-927-19 (1-279) x US-10-006-285-513 (1-3299)

```
Qy      9 IleTyrValGlyAsnLeuProGlyAspIleArgLysGluValGluAspLeuPheTyr  28
            ||||||||: ||||| ||||: : ||:::||||| |||||| ::|||||||
Db    203 ATCTACGTGGGTAACTTACCTCCAGACATCCGACAACCAAGGACATTGAGGACGTTCTAC  262

Qy     29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla  48
            ||: |||||| |: :|||| |||||||| ||||||||| ::|||
Db    263 AAATACGGCGCTATCCGACATCGACCTCAAGAATCGCCGGGGGACCGCCCTTCGCC  322

Qy     49 PheValGluPheGluAspProArgAspAlaAspArgAlaIleTyrGlyArgAspGlyTyr  68
            ||||||||||| |||||| ||:|||| |||| |||||||||||:::||||||||
Db    323 TTCGTTGAGTTCGAGGACCCGAGAGACGCGGAAGACGCGGTCGTATGGTCGCGACGGCTAT  382

Qy     69 AspPheAspGlyCysArgValGlyLeuArgValGluIleAlaHisGlyGlyArg-------  84
            ||||
Db    383 GATTACGATGGGTACCGTCTGCGGGTGGAGTTCCTCGAAGGCCGTGGAACAGGCCGA  442

Qy     85 ------------------ArgPheSerProSerValAsp  91
                              |||||
Db    443 GGCGCGGCGCGGGGTGGAGGTGGCGAGCTCCCCGAGGTCGCTATGGCCCC  493

Qy     92 ArgTyrSerSerSerTyrSerSerAlaSerArgAlaProSerArgArgSerAspTyrArgVal  111
            ::|||||||| ::|||
Db    494 CCATTCCAGGCGCGTCTGAAAAACAGAGTG  520

Qy    112 LeuValThrGlyLeuProProSerAlaSerTrpGlnAspLysLeuLysAspHisMetArgLys  131
            ||::|||||||: |||::
Db    521 GTTGTCTCTGGACTGCCTCCTCCAAGTGGAAGTTGGCAGGATTAAAGGATCACATGCGTGAA  580

Qy    132 AlaGlyAspValCysPheSerGluValPheProArgLysMetSerGlyValVal  151
            |||||:||:|||::
Db    581 GCAGGTGATGTATGTTTACCGAGATGGC------------------ACTGGTGTCGTG  631

Qy    152 AspTyrSerAsnTyrArgAspMetLysTyrAlaIleArgLysLeuArgAlaThrGluPhe  171
            :::|||| |||||||||| ::
Db    632 GAGTTTACGAACGAAAGAAGATATGACCTATGCAGTTCGAAAACTGGATAACACTAAGTT  691

Qy    172 Arg--AsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal  190
            :::  ||||
Db    692 AGATCTCATGAGGAGAAACTGCCTACATCCGGTTAAAGTTGATGGGCCGAAGTCCA  751

Qy    191 SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgGlyProSer  210
            |||
Db    752 AGTTATGGAAGATCTCGATCTCGAAGC---CGTAGTCGTAGCAGAAGCGTAGCAGAGC  808

Qy    211 CysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIle--SerProArg  229
            ::|||||
Db    809 ------------AACAGCAGGAGTCGCAGTTACTCCCCAAGGAGAAGCAGAGGATCACCACGC  859

Qy    230 SerArgProLeu---SerArgSerArgSer  238
            |||||
Db    860 TATTCTCCCCGTCATAGACAGATCTCGCTCT  889
```

RESULT 13
US-10-171-581-9
; Sequence 9, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14

```
Db    538 CCTCCCCGCCCTCCATGTTATTGCTTGTTGAGTTTGATAATGCTCGAGATGCCGAAGAT  597

Qy     61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla  80
            |||||| |||||||||| |||||:::|||
Db    598 GCAATTCGGCGGTCGAGATGATGACAACTTTGATGTTGTCGGTTAAGAGTGGAGCTTGCT  657

Qy     81 HisGlyGlyArgArgPheSerProSerValAspArgTyr-----------  93
            |||||||| ||| |||
Db    658 CATGGTGGTAGAGGGCCATCATCAGT--GATCGACGTGGATATGGAGGAGGAGGA  714

Qy     94 ------------SerSerTyrSerAlaSerArgAlaProSerArgArgSer  107
                        :::
Db    715 GGAGGAGGAGGTGGTGGCAGCGCGGCGGGGTCGTTTGGCATCTCACGCCATTCT  774

Qy    108 AspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLysLysAsp  127
            ||:::|| ::|||
Db    775 GAATTTCGAGTTATTGTTCGCGACTCCCATCTTCGCATCTTGGCAAGATTTGAAGGAT  834

Qy    128 HisMetArgLysAlaGlyAspValCysPheSerGluValPheProArgLysGlyMet  147
            :::
Db    835 CATATGCGAAAAGCTGGTGGTGATGTGTTTTCTGAAGTTCCCGTGATAGTGAAGGGACT  894

Qy    148 SerGlyValValAspTyrSerAsnTyrArgAspMetLysTyrAlaIleArgLysLeuAsp  167
            ||||||
Db    895 TTTGGCATTGTTGATTACACTAATTATGACGACATGAAGTATGCTATTCGGAAACTAGAT  954

Qy    168 AlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSer  187
            ::::||||
Db    955 GACACTGAGTTTAGAAATCCTGGGCGAGCTTATATTCGGGTGAGAGAAGATGAGAGT  1014

Qy    188 ArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArg  207
            |||||
Db   1015 AGT-----------------CGTTCTAGGAGCCCACGACAGA  1038

Qy    208 GlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIleSer  227
            |||||
Db   1039 AGTCCC------------------AGCAGGAGCAGAAGTCCAAAAAGA--------  1068

Qy    228 ProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerArgValSerArgSerGly  247
            ||||||
Db   1069 ----------------GTTAGAAGTAGATCTTTGGAGCGTTCTGTGTCTCTAGGTCTCGA  1110

Qy    248 SerLeuArgArgAlaGlyAlaAspTrpIleSerGlnSerArgSerArgSerArgSerSer  267
            |||
Db   1111 TCTAGATCGGCGATCAGTATCACCCATCAAATCT*CTAGG-----CCAAGATCACGTTCA  1164

Qy    268 ArgSerAsnSerProValSerProValIleSerGly  279
            |||||: ||||:|
Db   1165 CGATCTGGATCTCCT---CGTCAGGTGCGTGTCAGGT  1197
```

RESULT 12
US-10-006-285-513
; Sequence 513, Application US/10006285
; Publication No. US20030165854A1
; GENERAL INFORMATION:
; APPLICANT: Mary Jane Cunningham
; APPLICANT: Matthew R. Kaser
; TITLE OF INVENTION: MARKER GENES RESPONDING TO TREATMENT WITH TOXINS
; FILE REFERENCE: PA-0039 US
; CURRENT APPLICATION NUMBER: US/10/006,285
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 514
; SOFTWARE: PERL Program
SEQ ID NO 513
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030165854A1 411426.42
US-10-006-285-513

Alignment Scores:

```
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 9
;   LENGTH: 1717
;   TYPE: DNA
;   ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: M72709
;   DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-9

Alignment Scores:
Pred. No.:              1.56e-46        Length:         1717
Score:                  542.50          Matches:        136
Percent Similarity:     51.74%          Conservative:   28
Best Local Similarity:  42.90%          Mismatches:     45
Query Match:            37.39%          Indels:         108
DB:                                     Gaps:           9

US-10-014-927-19 (1-279) x US-10-171-581-9 (1-1717)

Qy    9  IleTyrValGlyAsnLeuProGlyProGlyAspIleArgLysCysGluValGluAspLeuPheTyr  28
            |||||||||    ||| :::|||:::||||||
Db  148  ATCTACGTGGGTAACTTACCTCCAGACATCCGAACCAAGGACATTGAGGACGTGTTCTAC  207

Qy   29  LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla  48
            :::|||||       |||||:::||||||
Db  208  AAATACGGGCGCTATCCGCGACATCGACCTCAAGATCGCCGCGGGGACCGCCCTTCGCC  267

Qy   49  PheValGluPheGluSerProArgArgAspAspAlaAlaIleTyrGlyArgAspGlyTyr  68
            |||||||||||:::| |||||||||||||||||
Db  268  TTCGTTGAGTTCGAGGACCCCGAGACGCAGAGGACGCGGGTGTATGGTCGCGACGGCTAT  327

Qy   69  AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg-----------  84
            ||||||
Db  328  GATTACCATGGGTACCGTCTGCGGGTGGAGTTCCTCGAAGCGGCCGTGGAACAGGCCGA  387

Qy   85  ----------------------ArgPheSerProSerValAsp  91
Db  388  GGCGGCGCGCGGGGGGTGGAGGTGGCGGAGCTCCCCGAGGTCCCTATCGGCCCC--------  438

Qy   92  ArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgArgAspTyrArgVal  111
            |||||||       :::|||     |||:::|
Db  439  AAGTACGGGCCGCCTATCCGCCAGAATCGGAAAAACAGAGTG  465

Qy  112  LeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLys  131
            :::|||||||||||
Db  466  GTTGTCTCTGGACTGGCTGCCTCCAAGTGAAGTTGGCAGGATTGGCAGGAATGCATGCGTGAA  525

Qy  132  AlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyValVal  151
            ||:::|||||:::|||:::|||:::||||||
Db  526  GCAGGTGATGTATGTTATGCTGATGTTTACCGAGATGGC-----------ACTGGTGTCGTG  576

Qy  152  AspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPhe  171
            :::::          |||
Db  577  GAGTTGTACGGAGAAGAAGATATGAACCTATGCAGTTCGAAAACTGGATAACACTAAGTTT  636

Qy  172  Arg-------------------------------------Asn-AlaPheSer--  176
            |||                                        ||:||
Db  637  AGATCTCATGAGGTAGGTTATACACGTATTCTTTCTTTGACCAGAATTGGATACAGTGG  696

Qy  176  ----------------------------------------------------  176
Db  697  TCTTAACAGTGGAATTTCAAGGTAAGGATTCAGGCAAGGTTGTCAAGTAAATTGCCAGAT  756

Qy  176  --------------------------------------------------  176
Db  757  TTCTGGTTTTTAGTTACATTGTATTACTTACGCATGTCTGAAGATGAGATGAAAGCTTAGAT  816

Qy  177  --------SerAlaTyrIleArgValArgGl  184
            :::|||||||||
Db  817  CTTTCAATGGAAAGTTCTGTCATCCAATAGGGAGAAACTGCCTACATCCCGGGTTAAAGT  876
```

(Continued second column)

```
Qy  184  uTyrGluSerArgSerValSerArgProAspSerLysSerTyrArgSerArgSe  204
            |||:::|||
Db  877  TGATGGGCCCAGAAGTTCCAAGTTATGGAAGATCTCGAAGC---CGTAGTCGTAG  933

Qy  204  rArgSerArgGlyProSerCysSerTyrSerArgSerLysSerValSerProAlaAr  224
            :::|||||||||||||||||||||
Db  934  CAGAAGCCGTAGCAGAGC-----------AACAGCAGGAGTCCGAGTTACTCCCCAAGGAG  984

Qy  224  gSerIle---SerProArgSerArgProLeu---SerArgSerArgSer  238
            |||||
Db  985  AAGCAGGAGATCACCACGCTATTCTCCCCGTCATAGCAGATCTCGCTCT  1033

RESULT 14
US-09-880-107-3342
; Sequence 3342, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;   APPLICANT: Horne, Darci T.
;   APPLICANT: Vockley, Joseph G.
;   APPLICANT: Scherf, Uwe
;   APPLICANT: Gene Logic, Inc.
;   TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;   FILE REFERENCE: 44921-5028-WO
;   CURRENT APPLICATION NUMBER: US/09/880,107
;   CURRENT FILING DATE: 2001-06-14
;   PRIOR APPLICATION NUMBER: US 60/211,379
;   PRIOR FILING DATE: 2000-06-14
;   PRIOR APPLICATION NUMBER: US 60/237,054
;   PRIOR FILING DATE: 2000-10-02
;   NUMBER OF SEQ ID NOS: 3950
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3342
;   LENGTH: 1069
;   TYPE: DNA
;   ORGANISM: Homo sapiens
; FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 U30825
US-09-880-107-3342

Alignment Scores:
Pred. No.:              3.77e-44        Length:         1069
Score:                  517.00          Matches:        117
Percent Similarity:     67.91%          Conservative:   29
Best Local Similarity:  54.42%          Mismatches:     42
Query Match:            35.63%          Indels:         28
DB:                                     Gaps:           6

US-09-880-107-3342 (1-279) x US-09-880-107-3342 (1-1069)

Qy    9  IleTyrValGlyAsnLeuProGlyGlyAspIleArgLysCysGluValGluAspLeuPheTyr  28
            ||||||||||||||      :::||| :::  ::::|||
Db   98  ATCTACGTGGGGGAACCTTCCGACCGACGTGCGGCGAGAAGGACTTGGAGGACCTGTTCTAC  157

Qy   29  LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGly-TyrAl  48
            |||||||||:::||||||  :::|
Db  158  AGTACGGGCCGCATCCGAGATCGAGATCGAGCTCAAGAA--CCGGACACGGCCTCGTGCCCTTCGC  216

Qy   48  aPheValGluPheGluSerProArgArgAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr  68
            ||||||||||:::|||:::|||
Db  217  CTTCGTCGTGCGCTTCGAGGACCCCGAGGATGCGCAGGATGCTATTATGGAAGGAAATGGTTA  276

Qy   68  rAspPheAspGlyCysArgLeuArgLeuArgGluVal----AlaHisGlyGlyArgPh  86
            :::||||||
Db  277  TGATTATGGCCAGTGTCGGCTTCGTGTGGAGTTCCCCAGGACTTATGGAGGTGGGGTGG  336

Qy   86  eSerProSerValAspArgTyrSerSerSerAlaSerArgAlaProSerArgArgAr  106
            ::::
Db  337  GTGGCCCCGTGGTGGGAGG-----------AATGGGCCTCCTACAAGAAG  375

Qy  106  gSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLy  126
            :::||||||
Db  376  ATCTGATTCCGAGTTCTTGTTTCAGGACTTCCTCCGTGTCCGTCAGGCAGCTGGCACCTGAA  435
```

```
Qy   86  esrProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgAr 106
              :::           |||                              :::  |||
Db  382  GTGGCCCGTGGTGGGAGG------------------AATGGGCCTCCTACAAGAAG  420

Qy  106  gSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLy 126
             |||:::||||||||||||||||||::::||||||||| |||  |||||||||||||||
Db  421  ATCTGATTTCCGAGTTCTTGTTCAGGACTTCCGTCAGGCAGCTCGGCAGGACCTGAA  480

Qy  126  sAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGl 146
             ||||||:::||||||||||||::||||||||||||||:::::||||||||||||||||
Db  481  GGATCACACATGGCGAGAAGCTGGGGGATGTCTGTTATGCTGATGTGCAGAAGGATGGAGTGGG 540

Qy  146  yMetSerGlyValValAspTyrSerAsnTyrArgAspMetLysTyrAlaIleArgLysLe 166
             ||||          :::|||||:::||||||||::::|||||||||||||||||||||
Db  541  GATG-------GTCGAGTATCTCAGAAAAGAAGACATGGAATATGCCCTGCGTAAACT  591

Qy  166  uAspAlaThrGluPheArg--AsnAlaPheSerSerAlaTyrIleArgValArgGluTy 185
             ||||||||||||||||          :::||||||||||||||||||||||||||
Db  592  GGATGACACCAAATTCCGTCTCCATGAGGGTGAAACTTCCTACACATCCGAGTT------- 643

Qy  185  rGluSerArgSerValSerArgSerProAspAspSerLysSerTyr--------ArgSe 202
             |||:::         |||       |||:::
Db  644  ----------------TATCCTGAGAGGAGCACCAGCTATGGCTACTCACGGTC  681

Qy  202  rArgSerArgSerArgGlyProSerCysSerTyrSerSerLys 216
             |||||||||||||||||||||  |||     |||
Db  682  TCGGTCTGGGTCAAGGGGCCGTGACTCTCCATACCAAAGCAGG 724
```

Search completed: March 26, 2004, 21:34:11
Job time : 441 secs

```
Qy  126  sAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGl 146
             ||||||:::||||||||||||::||||||||||||||:::||||||||||||||||
Db  436  GGATCACACATGGCGAGAAGCTGGGGGATGTCTGTTATGCTGATGTGCAGAAGGATGGAGTGGG 495

Qy  146  yMetSerGlyValValAspTyrSerAsnTyrArgAspMetLysTyrAlaIleArgLysLe 166
             ||||          :::|||||:::|||||||||::::|||||||||||||||||||||
Db  496  GATG-------GTCGAGTATCTCAGAAAAGAAGACATGGAATATGCCCTGCGTAAACT  546

Qy  166  uAspAlaThrGluPheArg--AsnAlaPheSerSerSerAlaTyrIleArgValArgGluTy 185
             ||||||          :::|||||||||||||||||||||||||||||
Db  547  GGATGACACCAAATTCCGTCTCATGAGGGTGAAACTTCCTACACATCCGAGTT------- 598

Qy  185  rGluSerArgSerValSerArgSerProAspAspSerLysSerTyr--------ArgSe 202
             |||:::         |||       |||:::
Db  599  -------------TATCCTGAGAGGAGCACCAGCTATGGCTACTCACGGTC  636

Qy  202  rArgSerArgSerArgGlyProSerCysSerTyrSerSerLys 216
             ||||||||||||||||||||||  |||     |||
Db  637  TCGGTCTGGGTCAAGGGGCCGTGACTCTCCATACCAAAGCAGG 679
```

RESULT 15
US-09-925-300-661
; Sequence 661, Application US/09925300
; Patent No. US20020151681A1
GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 661
; LENGTH: 1162
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1155)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-661

Alignment Scores:
Pred. No.:   4.2e-44   Length: 1162
Score: 517.00   Matches: 117
Percent Similarity: 67.91%   Conservative: 29
Best Local Similarity: 54.42%   Mismatches: 42
Query Match: 35.63%   Indels: 28
DB: 9   Gaps: 6

US-10-014-927-19 (1-279) x US-09-925-300-661 (1-1162)

```
Qy    9  IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
             |||||||||||||||||::||||||:::||:|||||||||||||||||||||||
Db  143  ATCTACGTGGGGAACCTTCCGACCGACGTGCCGAAGGACTTGGAGGACCTGTTCTAC  202

Qy   29  LysTyrGlyProIleValAspIleAspLeuLysIleAspProProArgProGly-TyrAl 48
             |||||||||||||  :::||||||||||:::|||||||  |||||||||  :::|
Db  203  AAGTACGGCCGCATCCCGACGATCGAGCTCAAGA--CCGGCACGGCCTTCGTGCCCTTCGC 261

Qy   48  aPheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTy 68
             |||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db  262  CTTCGTGCGCTTCGAGGACCCCGAGATGCAGAGGATGCTATTTATGGAAGAAATGGTTA 321

Qy   68  rAspPheAspGlyCysArgLeuValGluIle-----AlaHisGlyArgArgPh 86
             |||:::         |||              :::|||||||||||
Db  322  TGATTATGGCCAGTGTCGGCTTCGTGTGGAGTTCCCCAGGAGTCGCGGGTGG 381
```

```
                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:    March 26, 2004, 18:12:15 ; Search time 2829 Seconds
                                      (without alignments)
                                      2945.051 Million cell updates/sec

Title:          US-10-014-927-19
Perfect score:  1451
Sequence:       1 MSSRWNRTIVVGNLPGDIRK..........RSKSRSRSRSNSPVSPVISG 279

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2751 3289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:    55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO_spool/US10014927/runat_24032004_152926_9909/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10014927 @CGN_1_1_3437 @runat_24032004_152926_9909 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   EST:*
    1:  em_estba:*
    2:  em_esthum:*
    3:  em_estin:*
    4:  em_estmu:*
    5:  em_estov:*
    6:  em_estpl:*
    7:  em_estro:*
    8:  em_htc:*
    9:  gb_est1:*
   10:  gb_est2:*
   11:  gb_htc:*
   12:  gb_est3:*
   13:  gb_est4:*
   14:  gb_est5:*
   15:  em_estfun:*
   16:  em_estom:*
   17:  em_gss_hum:*
   18:  em_gss_inv:*
   19:  em_gss_pln:*
   20:  em_gss_vrt:*
   21:  em_gss_fun:*
   22:  em_gss_mam:*
   23:  em_gss_mus:*
   24:  em_gss_pro:*
   25:  em_gss_rod:*
   26:  em_gss_phg:*
   27:  em_gss_vrl:*
   28:  gb_gss1:*
   29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	992.5	68.4	604	9	AV826310	AV826310 AV826310
2	947	65.3	650	14	CB258427	CB258427 28-E01273
3	881	60.7	929	14	CK272077	CK272077 EST718155
4	880.5	60.7	948	14	CK270575	CK270575 EST716653
5	851	58.6	935	14	CK244849	CK244849 EST728486
6	845	58.2	1255	11	AY108591	AY108591 Zea mays
7	840.5	57.9	864	29	CG848083	CG848083 ZMMBBb031
8	839	57.8	877	14	CK267541	CK267541 EST713619
9	839	57.8	910	14	CK250556	CK250556 EST734193
10	839	57.8	921	14	CK256686	CK256686 EST740323
11	834.5	57.5	896	14	CK275124	CK275124 EST721202
12	829.5	57.2	874	29	CG848162	CG848162 ZMMBBb031
13	829.5	57.2	903	14	CK298330	CK298330 EST761044
14	829.5	57.2	930	14	CK285308	CK285308 EST748030
15	829	57.1	884	14	CK265450	CK265450 EST711528
16	827.5	57.0	779	13	BQ802976	BQ802976 WHE2832_C
17	822.5	56.7	833	14	CK273791	CK273791 EST719869
18	815.5	56.2	817	14	CK253962	CK253962 EST737599
19	814	56.1	839	14	CK252379	CK252379 EST736016
20	812	56.0	933	14	CK284417	CK284417 EST747139
21	811	55.9	933	14	CK269749	CK269749 EST715827
22	811	55.9	949	14	CK264971	CK264971 EST711049
23	809	55.8	897	14	CK298535	CK298535 EST761249
24	809	55.8	974	14	CK254408	CK254408 EST738045
25	808.5	55.7	1106	14	CK206059	CK206059 FGAS01763
26	807	55.6	989	14	CK270936	CK270936 EST717014
27	807	55.6	999	14	CK253886	CK253886 EST737523
28	805.5	55.5	806	12	BM109633	BM109633 EST735169
29	805	55.5	926	14	CK250645	CK250645 EST734282
30	804	55.4	809	14	CD879553	CD879553 AZO4.105K
31	803	55.3	919	14	CK259499	CK259499 EST705577
32	802	55.3	852	14	CD439565	CD439565 ELO1N0577
33	799	55.1	913	14	CK259500	CK259500 EST705578
34	799	55.1	924	14	CK264893	CK264893 EST710971
35	796.5	54.9	667	12	BG451868	BG451868 NF101A10D
36	796.5	54.9	812	14	CK272649	CK272649 EST718727
37	795	54.8	675	12	BG648263	BG648263 EST509882
38	795	54.8	840	14	CB650587	CB650587 OSJNEb15B
39	792	54.6	877	14	CK267276	CK267276 EST713354
40	792	54.6	900	14	CK262606	CK262606 EST708684
41	792	54.6	927	14	CK271717	CK271717 EST717795
42	789	54.4	627	12	BG135885	BG135885 EST468777
43	787	54.2	877	12	BG444501	BG444501 GA_Ea002
44	785.5	54.1	736	14	CA802008	CA802008 sau28e10.
45	784	54.0	647	12	BI931752	BI931752 EST551641

ALIGNMENTS

```
RESULT 1
AV826310
LOCUS       AV826310          604 bp  mRNA    linear   EST 01-APR-2002
DEFINITION  AV826310 RAFL8 Arabidopsis thaliana cDNA clone RAFL08-12-I05 5',
            mRNA sequence.
ACCESSION   AV826310
VERSION     AV826310.1  GI:19868370
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
REFERENCE    1  (bases 1 to 604)
AUTHORS      Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
             Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
             Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
             and Shinozaki,K.
TITLE        Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL      Unpublished (2002)
COMMENT      Contact: Motoaki Seki
             Plant Functional Genomics Research Group
             RIKEN Genomic Sciences Center
             3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
             Tel: 81-298-36-4359
             Fax: 81-298-36-9060
             Email: mseki@rtc.riken.go.jp
             An Arabidopsis full-length cDNA library was constructed essentially
             as reported previously (Seki et al., 1998).cDNA cleaved with BamHI
             and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
             al., submitted for publication) digested with BamHI and SalI. This
             clone is in a modified pBluescript vector. Please visit our web
             site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
             details.
FEATURES             Location/Qualifiers
     source          1..604
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL08-12-I05"
                     /dev_stage="rosette plants"
                     /lab_host="DH10B"
                     /clone_lib="RAFL8"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to
                     dehydration-treated (1, 2, 5, 10, 24 hr)"
ORIGIN

Alignment Scores:
Pred. No.:                  4.51e-63              Length:           604
Score:                      992.50                Matches:          197
Percent Similarity:         94.29%                Conservative:     1
Best Local Similarity:      93.81%                Mismatches:       1
Query Match:                68.40%                Indels:           11
DB:                         9                     Gaps:             1

US-10-014-927-19 (1-279) x AV826310 (1-604)

Qy    62  IleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHis    81
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3  ATTTATGGACGTGATGGTTATGATTTTGATGGGTGTCGACTTCGGGTTGAGATTGCACAT    62

Qy    82  GlyGlyArgArgPheSerGlyValAspArgTyrSerSerTyrSerAlaSerArg         101
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63  GGTGGTCGTAGATTTTCACCATCAGTTGATAGGTACAGCAGCTACAGTGCGAGCCGT      122

Qy   102  AlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSer   121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123  GCACCTTCAAGACGCTCTGACTACCACCGTGCTTGTGACCGGATTACCGCCTTCTGCTTCG  182

Qy   122  TrpGlnAspLeuLysAspHisMetArgLysMetLysAlaGlyAspValCysPheSerGluValPhe  141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183  TGGCAGGACCTTAAGGATCACATGCGCAAAGCTGGAGATGTCTGCTTCTCTGAAGTTTTC   242

Qy   142  ProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyr  161
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243  CCTGACCGTAAAGGCATGTCTGGGGTTGTGGATTATAGCAACTATGATGATGAAGTAC    302

Qy   162  AlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArg  181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303  GCAATAAGGAAACTTGATGCCACTGAATTTCGAAATGCTTTCTCTAGTGCTTATATACGG  362

Qy   182  ValArgGluTyrGluSerArgSerValSerArgSerProAspSerProLysSerTyrArg  201
          ||||||||||||||||||||||||||||||||||:|||||::::||||||||||
Db   363  GTGAGGGAATATGATGCGGAGGAGTGTCAGTCCGAAGCCCAGATGATTCTAAAAGCTATAGA  422

Qy   202  SerArgSerArgArgGlyProSerCysSerTyrSerCysSerLysSerArgSerValSer  221
```

```
Db   423  AGCAGGAGTCGGAGCCGTGGTCCAAGTGCTGTAGCTATAGTAGCAGAGAGCAGGAGTGTGTCA  482
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   222  ProAlaArgSerIleSerProArgSerArgSerProLeuSerArgArgSerGluTyrSer  241
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   483  CCTGCTAGATCCATTTCCCGCGTTCACGGCCCCTTAGTCTGTTCTCGCTCTATACAGC  542

Qy   242  SerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSer  261
          |||||||||||||||||||||||||||||||||||||||
Db   543  TCTGTCTCAAGG-------------------------TCCAATCAAGATCA  569

Qy   262  LysSerArgSerArgSerArgSerAsnSer  271
          |||||||||||||||||||||||||||
Db   570  AAATCAAGAACAAGATCAAGATCGAATTCT  599

RESULT 2
CB258427
CB258427       28-E012735-014-002-H16-T7R MPIZ-ADIS-014 Arabidopsis thaliana cDNA
LOCUS          CB258427               650 bp    mRNA    linear    EST 06-NOV-2003
DEFINITION     28-E012735-014-002-H16-T7R MPIZ-ADIS-014 Arabidopsis thaliana cDNA
               clone MPIZp771H16Q 5-PRIME, mRNA sequence.
ACCESSION      CB258427
VERSION        CB258427.1  GI:32883200
KEYWORDS       EST.
SOURCE         Arabidopsis thaliana (thale cress)
ORGANISM       Arabidopsis thaliana
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE      1  (bases 1 to 650)
AUTHORS        Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
               Mitchell-Olds,T. and Weisshaar,B.
TITLE          Large-scale identification and analysis of genome-wide
               single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL        Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE        22683290
PUBMED         12799357
COMMENT        Contact: Weisshaar B
               ADIS DNA core facility at MPIZ
               Max-Planck-Institute for Plant Breeding Research
               Carl-von-Linne Weg 10, 50829 Koeln, Germany
               Fax: 00492215062851
               Email: weisshaar@mpiz-koeln.mpg.de
               Insert Length: 650    Std Error: 0.00
               Plate: 2  row: H  column: 16
               Seq primer: T7R; CTAATACGACTCACTATAGGGA.
FEATURES             Location/Qualifiers
     source          1..650
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /cultivar="Cape Verde Island (Cvi)"
                     /db_xref="GABI:591109"
                     /db_xref="taxon:3702"
                     /clone="MPIZp771H16Q"
                     /tissue_type="whole plant"
                     /dev_stage="adult plant, mixed stresses"
                     /lab_host="E. coli TOP10"
                     /clone_lib="MPIZ-ADIS-014"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; cDNA
                     library from Arabidopsis thaliana, accession Cvi; six
                     weeks old total plants grown under long-day conditions in
                     soil, whole adult plants were treated for 24 hours with
                     different stresses; (1) at 4 Grad C in the dark, (2), at
                     37 Grad C in the dark, (3) lying in the lab after removing
                     from soil, (4) in the greenhouse after wounding leaves
                     with a forceps, (5) in the lab watering with a 150 mM
                     NaCl solution, (6) at 26 Grad C in the light/UV; equal
                     quantities of stressed plant material were pooled; library
                     was made at the Max-Planck-Institute for Plant Breeding
                     Research, Cologne, Germany; cloning sites SalI-NotI,
                     primer sites and orientation:
                     T7-SalI-CCACGGCTCG-5prime-cDNA-polyA-CC-NotI-SP6; Note:
                     Sequencing granted in the context of the GABI Arabidopsis
                     Verbund I: Genetic Diversity, 'Establishment of
```

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES   Location/Qualifiers
source   1..929
                  /organism="Solanum tuberosum"
                  /mol_type="mRNA"
                  /cultivar="Kennebec"
                  /db_xref="taxon:4113"
                  /clone="POAD303"
                  /tissue_type="abiotic stress treated leaf and root tissue"
                  /lab_host="DH10B-TonA"
                  /clone_lib="potato abiotic stress cDNA library"
                  /note="Vector: pCMVSport6.1; Site 1: EcoRI; Site 2: NotI;
                  supplier: Solanum tuberosum var. Kennebec plants were
                  grown from cuttings on a 16hr light/8 hr dark cycle at 25
                  C for 3-4 weeks. Abiotic stress conditions were applied to
                  four separate sets of plants. Set 1 involved saturation of
                  the soil with 150 mM NaCl and tissues were harvested at
                  following application of the salt stress (leaves: 2hr,
                  6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
                  Set 2 were grown under the standard conditions and then
                  were water stressed by withdrawal of further watering
                  application. Drought stressed plants were harvested after
                  cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
                  and 5d). Set 3 were grown under the standard conditions
                  and then were cold stressed by placement at 4 C. Cold
                  stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                  and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                  2d. Set 4 were grown under the standard conditions and
                  then were heat stressed by placement at 35 C. Heat
                  stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                  2d and 4d and heat-stressed roots were harvested at 6 hr,
                  12 hr, 1 d, and 4d. RNA was isolated from all tissues and
                  equal RNA from each tissue and stress was pooled to
                  construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:
Pred. No.:   929
Score: 1.23e-54   Length: 929
Percent Similarity: 76.41%   Matches: 194
Best Local Similarity: 68.31%   Conservative: 23
Query Match: 60.72%   Mismatches: 53
DB: 14   Indels: 14
                                              Gaps: 7

US-10-014-927-19 (1-279) x CK272077 (1-929)

Qy   4 ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyValAspIleArgLysCysGluVal 23
          :::::::::::::::::::::::::::::::::::   ::::   :::
Db   47 CGACTAAGTCGGACTATCTACGTCGGAAATCTTCCTCGTGATATTCGGGAGAGAAGTA 106

Qy   24 GluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysLysIleProProArg 43
          :::::::::::::::::::::::::::::::::: ::::::::::::::::
Db   107 GAAGATTTGTTTTACAAGTATGCCATTGTGGAAAATTGATTTGAAAGTTCCTCCTAGA 166

Qy   44 ProProGlyTyrAlaPheValGluPheGluAspProArgProArgAspAlaAspAlaIleTyr 63
          :::::::::::::::::::::::::::::   :::::::::::::::::::
Db   167 CCACCTGGTTATGCGTTTGAGTTTGAAGATCCTCGTGATGCTGATGCTATCTGC 226

Qy   64 GlyArgAspGlyTyrArgAspPheAspGlyCysArgLeuArgValGlyIleAlaHisGlyGly 83
          :::
Db   227 GGGCGTGATGGCTATGACTTTGACGGGCATCGTTGCCGAGTTGAACTGGTCACATGGTGG 286

Qy   84 ArgArgPheSerProSerValAspArgArgTyrSerSerTyrSerAlaSerArgAlaPro 103
          ::::::::   :::: :::::::::::::::
Db   287 CGAGGA---TCATCATCATATGATCGCCACAGTAGTTACAGTAGTGGGAGTCGTGGTGGA 343

Qy   104 --SerArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
          :::::: :::::::::::::::
Db   344 CTTTCTAGGCCGCTCGACTATCCGCTACCTGGTTCTCTGGACTACCATCTTCTGCTTCATG 403

RESULT 3
CK272077
LOCUS   CK272077   929 bp   mRNA   linear EST 12-DEC-2003
DEFINITION EST718155 potato abiotic stress cDNA library Solanum tuberosum cDNA
            clone POAD303 5' end, mRNA sequence.
ACCESSION   CK272077
VERSION   CK272077.1 GI:39829055
KEYWORDS   EST.
SOURCE   Solanum tuberosum (potato)
ORGANISM   Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1 (bases 1 to 929)
AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE   Generation of ESTs from abiotic stressed potato tissue
JOURNAL   Unpublished (2003)
COMMENT   Other_ESTs: EST718156
            Contact: Robin Buell

ORIGIN

high-efficiency SNP-based mapping tools and development of
methods for genome-wide mutation detection' PI: Bernd
Weisshaar Sequence submission managed by RZPD/GABI-Primary
database: http://gabi.rzpd.de. This clone is available
from RZPD; contact RZPD (clone@rzpd.de) for further
information."

Alignment Scores:
Pred. No.:   650
Score: 1.08e-59   Length: 650
Percent Similarity: 947.00   Matches: 177
Best Local Similarity: 98.33%   Conservative: 0
Query Match: 98.33%   Mismatches: 3
DB: 65.27%   Indels: 0
                            14   Gaps: 0

US-10-014-927-19 (1-279) x CB258427 (1-650)

Qy   1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyValAspIleArgLys 20
          :::
Db   111 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTGCCTGGAGATATTCGCAAG 170

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysLysIle 40
          :::
Db   171 TGTGAGGTTGAAGATCTCTTCTACAAGTATGGACCAATTGTGGACATTGATTTGAAGATT 230

Qy   41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgProArgAspAlaAspAlaAspAsp 60
          ::
Db   231 CCACCGAGACCCTCCTGGTTATGCGTTTGTCGAGTTGAAGATCCTCGTGATGCACACGAT 290

Qy   61 AlaIleTyrGlyArgAspGlyTyrArgAspPheAspGlyCysArgLeuArgValGlyIleAla 80
          ::
Db   291 GCAATTTATGGACGTGATGGTTATGATTTTGATGGGTGTCGACTTCGGTTGAGATTGCA 350

Qy   81 HisGlyGlyArgArgPheSerProSerValAspArgArgTyrSerSerTyrSerAlaSer 100
          ::
Db   351 CATGGTGGTCGTAGATTTTCACCATCAGGTGATAGGTACAGCAGCTACAGTGCGAGC 410

Qy   101 ArgAlaProSerArgArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
          ::
Db   411 CGTGCACCTTCTAGACGCTCCGACTATCCGCTACCTGGTGTGCTTGTGGACGGATTACCGCCTTCTGCT 470

Qy   121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
          ::
Db   471 TCGTGGCAGGAGACCTTAAGGATCACATGCGCAAAGCTGGAGATGTCTGCTTCTCTGAAGTT 530

Qy   141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerArgAsnTyrAspAspMetLys 160
          ::
Db   531 TTCCCTGACCGTAAAGGCATGTCTGGGGTTGGGGATTAATGCAACTATGATGATGAAG 590

Qy   161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
          ::
Db   591 TACGCAATAAGGAAACTTGATGCCACTGAATTTCGAAATCCTTTCTCTAGTGCTTATATA 650

```
Qy   123  GlnAspLeuLysAspHisMetArgLysAlaGlyPheSerGluValCysPheValPro  142
          |||||||||||||       ||||:::||||||||||::|||||||||||
Db   404  CAAGACTTGAAGGATCACATGCGACGAGCCTGGAGATGTGTGCTTCGTCTCCAAGTTTCCGA  463

Qy   143  AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrArgAspMetLysTyrAla  162
          |||||||||:::||||||||||:::|||:::||||||||||||||||
Db   464  GATCCGTGACGGTATGAGAGGGGATTGTGGACTATACCAACTATGATGATGAGATATGCG  523

Qy   163  IleAlaArgLeuAlaProAlaThrGluPheArgGluArgAsnAlaPheSerAlaTyrIleArgVal  182
          ||:::||||||  ::: ||||||||||||||:::||||||||
Db   524  ATAAAGAAACTTGATGACTCTGTTCCGCAAATCAATTCTCTCGAGCATATATTAGGGTG  583

Qy   183  ArgGluTyrGluSerArg--SerValSerArgSerProAspAsp------SerLysSer  199
          :::|||:::   ||||||||||:::|||||||:::
Db   584  GACAAGTATGATAAGAGGCCATAGCTATTCCAGGAGTCCAAGTCCATATTATTCAAGAAGC  643

Qy   200  TyrArgSerArgSerArgSerArgGlyTyrProSerCysSerTyrSerSerLysSerSer  219
          |||||| ||||||||||||   |::: |||||||:::|||
Db   644  ---AGAAGTTACTCAAGAAGCAGGAGTCCTCGACAAAGCTATAGCAGCCAGAGCGGAAGT  700

Qy   220  ValSerPro-----------AlaArgSerIleSerProArgSerArgPro  232
          :::   ||:::||||||||||||
Db   701  GTATCTCCTAGGGGTAAAATACTGTCGCTCTGTGTCTGTCTCACCC--TCAAGGGCT  757

Qy   233  LeuSerArgSerArgSerLeuTyrSerArgSerValSerArgSerGlyLysSerLeuArgGlyAla  252
          :::|||||||||||||||||:::|||||||||||||
Db   758  TTTTCCCCTGCTCGCTCTCTTCCAGATCTCATCAAGATCCAGATCCCACTTCACCT  817

Qy   253  GlyAspTrpIleSerGlnSerArgSerArgSerArgSerArgSerArgSerArgSerAsnSerPro  272
          ||||||     |||   |||||||||
Db   818  CCGCCTCGTTGGAGCGCACGGAAGGAGCCCGAGCAGGAGAGTCTGTCCTATTCTCGC  877

Qy   273  ValSerProVal  276
          ||| |::
Db   878  TCTTCGGGGATA  889
```

```
RESULT 4
CK270575
LOCUS        CK270575            948 bp    mRNA    linear   EST 12-DEC-2003
DEFINITION   EST716653 potato abiotic stress cDNA library Solanum tuberosum cDNA
             clone POACU05 5' end, mRNA sequence.
ACCESSION    CK270575
VERSION      CK270575.1  GI:39827553
KEYWORDS     EST.
SOURCE       Solanum tuberosum (potato)
  ORGANISM   Solanum tuberosum
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE    1  (bases 1 to 948)
  AUTHORS    Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
  TITLE      Generation of ESTs from abiotic stressed potato tissue
  JOURNAL    Unpublished (2003)
COMMENT      Other_ESTs: EST716654
             Contact: Robin Buell
             The Institute for Genomic Research
             9712 Medical Center Dr, Rockville, MD 20850, USA
             Email: potato-array@tigr.org
             Clones can be requested from TIGR via potato@tigr.org
             Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1..948
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POACU05"
                     /tissue_type="abiotic stress treated leaf and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="potato abiotic stress cDNA library"
                     /note="Vector: pCWSport6.1; Site 1: EcoRI; Site 2: NotI;
                     supplier: Solanum tuberosum var. Kennebec plants were
```

grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
application. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

```
ORIGIN

Alignment Scores:
Pred. No.:                      948
Score:                 1.38e-54
Percent Similarity:    880.50     Matches:          188
Best Local Similarity: 75.72%     Conservative:      21
Query Match:           68.12%     Mismatches:        44
DB:                    60.68%     Indels:            23
                       14         Gaps:               6


US-10-014-927-19 (1-279) x CK270575 (1-948)

Qy     4  ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluVal   23
          :::|||||||||||||||||:::|||||||||||||||:::
Db    62  CGTCTAAGTCGGACTATCTACGTCGGAAATCTTCCGTGGTGATATTCGGGAGAGAAGTA  121

Qy    24  GluAspLeuPheTyrLysTyrGlyTyrGlyProIleValGlyIleLeuAspLeuLysIleProProArg  43
          |||||||||||||||:::|||||||:::||||||||||||||||
Db   122  GAAGATTTGTTTTACAAGTATGGTCCCATTGTGGAAATTGATTTGAAAGTTCACCTAGA  181

Qy    44  ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAspAlaIleTyr   63
          |||||||||:::|||||||||||||||||||||||:::
Db   182  CCACCTGGTTATGCGTTCGTAGAGTTTGAAGATCCTCGTGATGCTGATGCCATCCGT  241

Qy    64  GlyArgAspGlyTyrArgAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGly   83
          |||||||:::|||||||||||:::|||||||||||||||||||
Db   242  GGGCGTGATGGCTATGACTTTGATGGGCATCGTCTGCGAGTTGAACTTGCACATGGTGGG  301

Qy    84  ArgArgPheSerProSerValAspArgArgTyrSerSerSerTyrSerAlaSerArgAlaPro  103
          :::||||||:::||||||::|||||||||||||:::
Db   302  CGAGGA--TCATCATCATGATCGCCACAGTAGTTACAGTAGTGGGAGTTCGTGGTGAA  358

Qy   104  --SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp  122
          |||||||||||:::|||||||||||:::||||||
Db   359  TTTTCTAGGCGCCTGACTATCGCGTACTGGTCTCCTGACTACCATCTTCTGTTCATGG  418

Qy   123  GlnArgLysLysAspHisMetArgLysAlaGlyPheSerGluValCysPheGluValPhePro  142
          |||||||||||||::::|||||||||||||||::||||||||
Db   419  CAAGACTTGAAGGATCACATGCGACGAGCCTGGAGATGTGTGCTTCGTCTCCAAGTTTCCGA  478

Qy   143  AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrArgAspMetLysTyrAla  162
          |||||||||:::||||||||||:::|||:::||||||||||||||||
Db   479  GATCCGTGACGGTATGAGAGGGGATTGTGGACTATACCAACTATGATGATGAGATATGCG  538

Qy   163  IleAlaArgLeuAspAlaThrGluPheArgGluArgAsnAlaPheSerAlaTyrIleArgVal  182
          ||:::||||||  ::: |||||||||||||||||
Db   539  ATAAAGAAACTTGATGCTGTTCCGCAATCAATTCTCTCGAGCATATATTAGGGTG  598

Qy   183  ArgGluTyrGluSerArg--SerValSerArgSerProAspAsp------SerLysSer  199
          :::|||:::   ||||||||||:::|||||||:::
Db   599  GACAAGTATGATAAGAGGCCATAGCTATTCCAGGAGTCCAAGTCCATATTATTCAAGAAGC  658
```

RESULT 5

```
Qy   200 TyrArgSerArgSerArgArgGlyProSerCysSerTyrSerSerLysSerArgSer 219
             |||||     ||||||||||    |||
Db   659 ---AGAAGTTACTCAAGAGCAGCAGAGTCCTCGACGAAGCTATAGCAGCCAGAGCGCGGAAGT 715

Qy   220 ValSerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeu 239
                      :::|||                                ||:
Db   716 GTATCTCCTAGGGGTAAATACTCCGTCCGTCTGTCGTCTCGTCTCGTGTCTCCACCTTCAAGGAT--- 772

Qy   240 TyrSerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSer 259
                                                                    :::
Db   773 -------------------------------------------------------TTTTCCCCTGCT 784

Qy   260 ArgSerLysSerArgArgSerArgSerAsnSerProValSerPro 275
             ||||||||||    ||||||||||    ||||||||
Db   785 CGCTCTCATTCAGCCAGGATCTCTATCAAGATCCAGATCTCCACTTTCACCT 832
```

RESULT 5

```
CK244849
LOCUS       CK244849              935 bp    mRNA        linear      EST 12-DEC-2003
DEFINITION  EST728486 potato callus cDNA library, normalized and full-length
            Solanum tuberosum cDNA clone POCAA66 5' end, mRNA sequence.
ACCESSION   CK244849
VERSION     CK244849.1  GI:39790831
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
ORGANISM    Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 935)
AUTHORS     Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE       Generation of ESTs from potato callus tissue
JOURNAL     Unpublished (2003)
COMMENT     Other ESTs: EST728485 EST728487
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1..935
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POCAA66"
                     /tissue_type="callus"
                     /lab_host="DH10B-TonA"
                     /clone_lib="potato callus cDNA library, normalized and
                     full-length"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: RNA was isolated from Solanum tuberosum var.
                     Kennebec callus tissue grown on solid media."
ORIGIN
```

Alignment Scores:

```
Pred. No.:                 1.97e-52          Length:          935
Score:                     851.00            Matches:         183
Percent Similarity:        75.19%            Conservative:    20
Best Local Similarity:     67.78%            Mismatches:      51
Query Match:               58.65%            Indels:          16
DB:                        14                Gaps:            6
```

```
US-10-014-927-19 (1-279) x CK244849 (1-935)

Qy     4 ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluVal 23
             |:|||           |||        |||       ||||||:|
Db    13 CGTCTAAGTCGGACTATCTACGTCCGAAATCTTCCTGGTGATATTCGGCAGAGAAGTA 72

Qy    24 GluAspLeuPheTyrLysPheTyrGlyProIleValAspIleAspIleLysLeuLysAspIleLeuProProArg 43
             ||||||||||||||||||              :::||||||||
Db    73 GAAGATTGGTTTACAAGTACATGTCCCATTGTGAAATTGATTTGAAAGTTCCACCTAGA 132
```

RESULT 6

```
AY108591
LOCUS       AY108591              1255 bp    mRNA        linear      HTC 17-OCT-2002
DEFINITION  Zea mays PCO125156 mRNA sequence.
ACCESSION   AY108591
VERSION     AY108591.1  GI:21211708
KEYWORDS    HTC.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1255)
AUTHORS     Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE       Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
JOURNAL     Unpublished (2002)
REFERENCE   2  (bases 1 to 1255)
AUTHORS     Coe,E.H.
TITLE       Direct Submission
JOURNAL     Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
```

```
Qy    44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAspAlaIleTyr 63
             |||||||||||     ||||||||||||    ||||||||||||||||||||
Db   133 CCACCTGGTTATCGCGTTCGTAGAGTTTGAAGATCCTCGTGATGCTGATGATGCCATCCGT 192

Qy    64 GlyArgArgAspGlyTyrAspPheAspGlyCysArgArgLeuArgValGluIleAlaHisGlyGly 83
             :::||||||||                    ||:||||
Db   193 GGGCGGTGATGGCTATGACTTTGATGGGCATGCGTCTTGCGAGTTGAACTTGCACATGGTGGG 252

Qy    84 ArgArgPheSerProSerValAspArgTyrSerSerArgSerArgAlaSerArgAlaPro 103
             ||||     ||:|||     |:|||
Db   253 CGAGGA---TCATCATCATATGATCGCCACAGTAGTTACAGTAGTGGGAGTCGTGGTGAA 309

Qy   104 --SerArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
             ||||||||||    ||||||||||||||||||||||||
Db   310 TTTTCTAGGCGCTCTGACTATCGCGTACTCGGTCTCTCGACTACCATCTTCGCTTCATGG 369

Qy   123 GlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142
             |||||    ||||||||    ||:||||||||    ||||||||    :::||
Db   370 CAAGACTTGAGGATCACATGCGACCAGCTGGAGATGTCTGCTTTCCTCAAGTTTTCCGA 429

Qy   143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAla 162
             :::||||||              ||||||||    ||||:||:||:||
Db   430 GATCGTGATGGTATGAGAGGGATTGTGGACTATACCAACTATGATGATGATGATATGCG 489

Qy   163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
             :::|:||:||        |||         ||::||        |:||||
Db   490 ATAAAGAAACTTGATGACTCTCTGTTTCGCAATCAATTCTCCGAGCATATATTAGGGTG 549

Qy   183 ArgGluTyrGluSerArg-----SerValSerArgSerProAspAsp------SerLysSer 199
             ::|:|||                 ||||||||||||||||
Db   550 GACAAGTATGATAAGAGGCATAGTCTATTCCAGGAGTCCAAGTCCATATTATTCAAGAAGC 609

Qy   200 TyrArgSerArgSerArgArgGlyProSerCysSerTyrSerSerLysSerArgSer 219
             |||||     |||||||||    |||
Db   610 ---AGAAGTTACTCAAGAGCAGCAGAGTCCTCGACGAAGCTATAGCAGCCAGAGCGCGGAAGT 666

Qy   220 ValSerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeu 239
                      :::|||                          :::|||
Db   667 GTATCTCCTAGGGGTAAATACTCCGTCCGTCTGTCGTCTCGTCTCGTGTCTCCACCTTCAAGGGATTTT 726

Qy   240 TyrSerSerValSerArgSerGlySerLeuLeuArgAla----------- 252
             :::||
Db   727 TCCCCTGCTCGCTCTCATTCAGCCAGGATCTCTATCAAGATCCAGATCTGGCTTCCGCGGCGATCTGGAATTGCTATATAG 786

Qy   253 ----------GlyAspTrpIleSerGlnSer 259
             ||:||||
Db   787 ACAGGTGATGGTTAACGTGTGCTCAAATTCA 816
```

```
Db    696  AAACGTGGTCGCTCCTACTCACGTAGCCGAAGCCCAAGTCGTAGTTAC---AGCAAAGC  752
           |||      ||||||||||||||||||||||||||||||||||||||||||      |||||||||
Qy    218  ArgSerValSerProAlaArgSerIleSerProArgSerArg-------ProLeuSer  234
                                                       ::::
Db    753  AGGAGTCCGAGTAA------------TCACCCAGGACTGGCCGTTCATCATCTAGATCC  800

Qy    235  ArgSerArgSerLeuTyrSerSerValSerArgGlySerLeuLeuArgAlaGlyAsp  254
           |||||||||||||||||||      :::::
Db    801  CGGTCAAGATCTGTT---TCTTCGTTCAAGGTCCCCA----------------------  836

Qy    255  TrpIleSerGlnSerArgLysSerArgSerArgSerArgSerAsnSerPro---Val  273
                 |||||  |||      ||||||      ::::::|||
Db    837  ------TCAAAGGACGTTCTCCATCAGATCACCAGCAAGATCGAAATCTCCTAATGTT  890

Qy    274  SerPro  275
           |||||||
Db    891  TCTCCA  896
```

```
RESULT 7
CG848083
LOCUS       CG848083                      864 bp      DNA      linear   GSS 18-NOV-2003
DEFINITION  ZMMBBo0319124.r ZMMBBb Zea mays subsp. mays genomic clone
            ZMMBBo0319124 3', genomic survey sequence.
ACCESSION   CG848083
VERSION     CG848083.1  GI:3837944
KEYWORDS    GSS.
SOURCE      Zea mays subsp. mays (maize)
  ORGANISM  Zea mays subsp. mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 864)
  AUTHORS   Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
            and Wing,R.
  TITLE     Sequencing of the maize genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: T7
            BACKWARD: M13r
            Plate: 0319  row: I   column: 24
            Seq primer: M13r
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..864
                     /organism="Zea mays subsp. mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /sub_species="mays"
                     /db_xref="taxon:4578"
                     /clone="ZMMBBo0319124"
                     /lab_host="DH10B"
                     /clone_lib="ZMMBBb"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Zea mays L. ssp. mays"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.04e-51               Length:          864
Score:                  840.50                 Matches:         175
Percent Similarity:     79.20%                 Conservative:    23
Best Local Similarity:  70.00%                 Mismatches:      37
Query Match:            57.93%                 Indels:          15
DB:                     29                     Gaps:            5
```

```
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
```

```
FEATURES             Location/Qualifiers
     source          1..1255
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:637293"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
```

```
ORIGIN
Alignment Scores:
Pred. No.:              8.19e-52               Length:          1255
Score:                  845.00                 Matches:         186
Percent Similarity:     73.05%                 Conservative:    20
Best Local Similarity:  65.96%                 Mismatches:      44
Query Match:            58.24%                 Indels:          32
DB:                                            Gaps:            8
```

```
US-10-014-927-19 (1-279) x AY108591 (1-1255)

Qy      1  MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys  20
           |||||                         :::    :::
Db    126  ATGAGCAGGCGGTGGAGCCGCACGATCACGTCGGGAACCTCCCCGGCGACATCCGGGAG  185

Qy     21  CysGluValGluAspLeuPheTyrLysTyrGlyTyrProIleValAspIleAspLeuLysIle  40
           |||         |||||                 |||
Db    186  AGGGAGGTGGGAGGATCTGTTCTCACAAGTATGGTAAATTGTTGACATTGACCTGAAGGTC  245

Qy     41  ProProArgProGlyTyrAlaPheValGluPheGluAspProAlaArgAlaAspAsp  60
                     :::           :::     :::::
Db    246  CCCCCAAGACCACCACCTGGTTATGCTTTGTTGTTGAGGATCCTCGTGATGCTGAGGAG  305

Qy     61  AlaIleTyrArgGlyArgAspGlyTyrAspPheArgGlyCysArgLeuArgValGluIleAla  80
           |||||        ::|||                 :::
Db    306  GCAATTGCTGGACGCGGATGGATCAACTTTGATGACACCGGTCTAAGAGTGGAGGCTGCT  365

Qy     81  HisGlyGlyArgArgPheSerProSerValAspArg--------TyrSerSerSerTyr  97
                :::                                   :::
Db    366  CATGGTGGTAGAGGTTAAATGCTTCCTCCGCATGAGTCGCTTCAAGTGGCTTTGGTCGGGTGGT  425

Qy     98  SerAlaSerArgAlaProSerArgArgArgSerAspTyrArgValLeuValThrGlyLeuPro  117
                      :::                 :::
Db    426  GGAGCACGTCGTGGTGTGTCGAGACAACTCAGAGTATCCGTGTTCCTTGTCACTGGACTGCCT  485

Qy    118  ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysValLysAspValCysPhe  137
           |||::::    :::    |||||
Db    486  TCTTCTGCATCATGCAGGAGATTTAAAGGATCATATGCGGAAGGCTGGTGATGTTTGTTC  545

Qy    138  SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp  157
           ::::       ::::::
Db    546  TCTGAAGTGTATCCGGAAGCGGGTGGTGCACCGTAGGAATTGTGGACTACACAAATTATGAT  605

Qy    158  AspMetLysTyrArgAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer  177
           |||:::                              ||||||
Db    606  GATATGAAATATGCTATAAAGAGCTGATGATACTGAATTCAGGAACGCCTTTGGGCGGA  665

Qy    178  AlaTyrIleArgValArgGluTyrGluSerArgValSerArgSerProAspAspSer  197
                   ::                                    :::
Db    666  GCCTATATAAGGGTGAAGGAATAT--------------------------AACGGC  695

Qy    198  LysSerTyrArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSer  217
```

US-10-014-927-19 (1-279) x CG848083 (1-864)

```
Qy      1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys  20
          ||| ||| ||| |||     || |::::||| ||||    |:::||
Db    103 ATGAGTAGCCGTTCCAGCAGGACTCTTTACGTTGGAAATCTTCCGGTGATGTTCCGCGAG 162

Qy     21 CysGluValGlyLeuAspLeuPheTyrLysTyrGlyProIleValAspIleAspLysIle  40
          ||||| ||| ||:||   ||||| || ||||||  |: | |   ||| || | |
Db    163 AGGGAAGTGGAAGATTGTTTTATAAGTAAGTATGGTCCTATAGCTCAAATTGACTTGAAGATT 222

Qy     41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp  60
          ||  | ||| ||||| || || || || ||: ||:   |::||
Db    223 CCACCAAGGCCTCCAGGTTATGCCATTGTTGGAGTTTGAGAGAGCTCGAGATGCTGAAGAT 282

Qy     61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla  80
          ||| ||| ||: ||| ||| | || |||:::::| |   |: | | |
Db    283 GCCATTCGTGGTCGTGATGGAATGACTTTGAGGGGCATCGTTTACGGGTGAACTTGCT 342

Qy     81 HisGlyGlyArgPheSerProSerValAspArgTyrSerSer---SerTyrSerAla  99
          ||| || | |    || || ||| | |:::        ||||
Db    343 CATGGTGGTCGTGGGCGCTCA---TCCATAGATCGTCACAGCAGTTATAGTAGTGGGCGT 399

Qy    100 SerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSer 119
          || ||| | || |||| | ||: ||| ||:::| |  ||: || | ||| |||  ||
Db    400 GGACGTGGACCATCCAGGCGTTCTGAAATCGGCTGCTAGTTACTGGATTGCCATCTTCT 459

Qy    120 AlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyTyrAspValCysPheSerGlu 139
          ||| || || |  ||: |:|   || | ||:::| |  | |   ||:::|::
Db    460 GCTTCACTGGCAGGACCTCAAGGATCACATGCGTCGAGCCAGGAGATGTTTGTTTTCCAA 519

Qy    140 ValPheProAspArgLysGlyMetSerGlyValValAlaAspTyrSerAsnTyrAspAspMet 159
          || || ||| ||| |:::||||        ||| ||:::|::|| | |
Db    520 GTTTTCCCGATGCGTAGTGGGACTAGCAGGGATTGTGGACTACACCAACTATGATGATATG 579

Qy    160 LysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyr 179
          |||| ||| ||||  |  ::: ||::|:::|||| ||| |
Db    580 AAGTATGCTATTAAGAAACTGGATGACTCTGAGTTTCGTAATGCATTTTCTCGGGCATAT 639

Qy    180 IleArgValArgGluTyrGluSer------------------------Arg 188
          ::: ||| |||:::| | ||:::
Db    640 GTTCGGGTTAAGGAATATGATTCTAGGCGGGATTCCTCTGGAAGCCCTAGTCGTGGCCGA 699

Qy    189 SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgArgGly 208
          || ||| || || || || |::::|||  ||| ||| || ||
Db    700 TCTCTCTCAAGAGCAGAAGCCAAGCCGAAGT---AGAAGCCGAGCAGAAGCCGTGGT 756

Qy    209 ProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIleSerPro 228
          ||| | :::||||||||
Db    757 CGGAGCTATAGC---CGAAGCAAGAGCCGCAGCAAGTCTCCAAAGGCAAAGCCTTCGCGC 813

Qy    229 ArgSerArgProLeuSerArgSerArgSer 238
          ::: | |
Db    814 AAGTCACCTGAAAAGTCTAGATCAAGGTCT 843
```

RESULT 8
CK267541

```
LOCUS       CK267541                 877 bp    mRNA    linear    EST 12-DEC-2003
DEFINITION  BST713619 potato abiotic stress cDNA library Solanum tuberosum cDNA
            clone POACB28 5' end, mRNA sequence.
ACCESSION   CK267541
VERSION     CK267541.1  GI:39824519
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
ORGANISM    Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
  TITLE     Generation of ESTs from abiotic stressed potato tissue
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
```

9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.

```
FEATURES             Location/Qualifiers
     source          1..877
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POACB28"
                     /tissue_type="abiotic stress treated leaf and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="potato abiotic stress cDNA library"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: Solanum tuberosum var. Kennebec plants were
                     grown from cuttings on a 16hr light/8 hr dark cycle at 25
                     C for 3-4 weeks. Abiotic stress conditions were applied to
                     four separate sets of plants. Set 1 involved saturation of
                     the soil with 150 mM NaCl and tissues were harvested at
                     following application of the salt stress (leaves: 2hr,
                     6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
                     Set 2 were grown under the standard conditions and then
                     were water stressed by withdrawal of further watering
                     applications. Drought stressed plants were harvested after
                     cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
                     and 5d). Set 3 were grown under the standard conditions
                     and then cold stressed by placement at 4 C. Cold
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d. Set 4 were grown under the standard conditions and
                     then were heat stressed by placement at 35 C. Heat
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d and 4d and heat-stressed roots were harvested at 6 hr,
                     12 hr, 1 d, and 4d. RNA was isolated from all tissues and
                     equal RNA from each tissue and stress was pooled to
                     construct the cDNA library. RNA sample."
```

ORIGIN

```
Alignment Scores:
Pred. No.:                                    877
Score:                1.37e-51       Length:        185
Percent Similarity:   839.00         Matches:        22
Best Local Similarity: 72.63%        Conservative:   48
Query Match:          64.91%         Mismatches:     30
DB:                   57.82%         Indels:          7
                      14             Gaps:
```

US-10-014-927-19 (1-279) x CK267541 (1-877)

```
Qy      3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
          ||| ||:::| ||||| |:::|| |:::||||||||||||::: ||||
Db     23 AGTCGGTCCGAGTAGGACACTTTATGTTGGCAATCTTCCTGGTGATGTTCGTGAGCGTGAA 82

Qy     23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLysIleProPro 42
          ||:| |  ||||||||||||||| |||| |:: || ||| |||||:::
Db     83 GTTGGAAGATCTGTTTCAAGTATCGGTCCGATAGTGGCCGATATTGAGATGCTGAAAATTCCACCA 142

Qy     43 ArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAspAlaIle 62
          |||||| ||| || |||||| | || |||:::||||:::||:::
Db    143 AGACCCCTGGTTATGCTTTTGTTGAGTTTGAAGAGGCTCGACGATGCTGAAGATGCTATT 202

Qy     63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
          ||| || ||| || |:::||||| |:::||| ||| |:::
Db    203 CGTGGTCGTGATGGCTATGATTTTGATGGGCACCGTCTGCGGGTTGAGCTTGCACATGGC 262

Qy     83 GlyArgArgPheSerProSerValAspSerArgTyrSerSerAla--------- 99
          || ||:::|| || ||| ||:|:::|||| || || ||
Db    263 GGGCGGTGGTAACTCATCATCAAATGATCGTTATGGTGGCGGCGGGTGGCGGCGGC 322

Qy    100 ------SerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGly 115
                :::|| ||| | |||:::|::| || |:::|| |:::
Db    323 CGTGGTCAGCGTGGTGGAGGAGTGTCCAGACGTTCGATTTTCGAGTTGTTAGTTACAGGA 382
```

```
ORIGIN              Kennebec callus tissue grown on solid media."

Alignment Scores:
Pred. No.:
Score:                  1.44e-51      Length:        910
                        839.00        Matches:       185
Percent Similarity:     72.63%        Conservative:  22
Best Local Similarity:  64.91%        Mismatches:    48
Query Match:            57.82%        Indels:        30
DB:                     14            Gaps:          7

US-10-014-927-19 (1-279) x CK250556 (1-910)

Qy    3  SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu   22
              ::  |||  :::||  |  ::|:  :|::::| :|||||||||||||
Db   23  AGTCCGGTCGGAGTAGGACACCTTATGTTGGCAATCTTCCTGGTCGATGTTCGTGAGCGTGAA  82

Qy   23  ValGluAspLeuPheTyrLysTyrGlyTyrProIleValAspIleAspIleAspLeuLysIleProPro  42
              ||||||||||||||
Db   83  GTGGAAGATCTGTTTCACAGTAGGCCCGATAGCTCATATTGAGCTGAAAATTCCACCA   142

Qy   43  ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAspAlaAlaIle   62
              :::||:::|||:::
Db  143  AGACCCCCTGGTTATGCTTTTGTTGAGTTTGAGGAGGCTCGACGATGCTATT   202

Qy   63  TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly   82
              ||||||:::|||:::|||
Db  203  CGTGGTCGTGATGGCTATGATTTTGATGGGCACCCGTCTGCGGGTTGCACATGGC   262

Qy   83  GlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerSerAla----------   99
              ||||||
Db  263  GGGCGGCGTGGTAACTCATCATCACAAATGATCGTTATGGTGCGGCGCGGTGGTGGCGGCGGC  322

Qy  100  ------SerArgAlaProSerArgSerArgArgSerArgAspTyrArgValLeuValThrGly  115
              ||||||
Db  323  CGTGGTCAGCGTGCGTGGTGGAGGAGTGTCCAGACGTTCTGATTTTCGAGTGTTAGTTACAGGA 382

Qy  116  LeuProProSerAlaSerTrpGlnAspLeuLeuLysAspHisMetArgLysAlaGlyAspVal 135
              ||||||||||||||
Db  383  TTACCCCATTCAGCATCGTGGCAGGATCTCAAGGATCACATGCGTCGTGCTGGGGGATGTT   442

Qy  136  CysPheSerGluValPheProArgLysGlyMetSerGlyValValAspTyrSerSerAsn   155
              :::||:::||:::
Db  443  TGTTTCTCCCAAGTTTTCCGTGATGGGAGTGGCACCACTGGTATAATTGATTATACAAAC   502

Qy  156  TyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPhe   175
              ||||||:::|||:::|||
Db  503  TATGATGACATGAAATATGCTATCAAAAAGCTTGATGACTCTGAGTTTCGGAATGCCTTT   562

Qy  176  SerSerAlaTyrIleArgValArgGluTyrGlu--SerArgSerValSerArgSerPro   194
              ||:::|||
Db  563  TCTTCTGCGACAATTCAGTTAAGGAATATGACCCGCAGTCGAGCCGCAGCCGC   622

Qy  195  AspAspSerLysSerTyr---------------ArgSerArgSerArg   207
              |||:::|||
Db  623  AGTCGGAGCCGGTTCTTACTTCCAAAGGAAGAGTGTTAGTCGAAGCCGTAGCCGA   682

Qy  208  GlyProSerCysSerTyrSerSerArgSerValSerProAlaArgSerIleSer   227
              |||
Db  683  AGTAGAAGT---------CGTAGCAAGAGCAGAAGCAAATCCTAAAGTAAGTCTTCA   733

Qy  228  ProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSerGly   247
              :::|||
Db  734  AAGGCGTCAAGA------TCTCGTCGTCCTG---TCTTCTCCGTCGTTCTGGG   784

Qy  248  SerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgLysSerArgSerArgSer   267
              ||:::|||
Db  785  ----------------------TCAAAAGGACGTCCTGTCTCAAGATCTCCATCA   817

Qy  268  ArgSerAsnSerPro  272
              ||||||
Db  818  AGGTCCAGATCGCCA  832
```

```
RESULT 9
CK250556
LOCUS       CK250556      910 bp    mRNA    linear   EST 12-DEC-2003
DEFINITION  EST734193 potato callus cDNA library, normalized and full-length
            Solanum tuberosum cDNA clone POCBG90 5' end, mRNA sequence.
ACCESSION   CK250556
VERSION     CK250556.1  GI:39802742
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
ORGANISM    Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiid; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 910)
AUTHORS     Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE       Generation of ESTs from potato callus tissue
JOURNAL     Unpublished (2003)
COMMENT     Other_ESTs: EST734194
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1..910
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POCBG90"
                     /tissue_type="callus"
                     /lab_host="DH10B-TonA"
                     /clone_lib="potato callus cDNA library, normalized and
                     full-length"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: RNA was isolated from Solanum tuberosum var.
```

```
RESULT 10
LOCUS       CK256686                921 bp    mRNA    linear   EST 12-DEC-2003
DEFINITION  EST740323 potato callus cDNA library, normalized and full-length
            Solanum tuberosum cDNA clone POCCV33 5' end, mRNA sequence.
ACCESSION   CK256686
VERSION     CK256686.1  GI:39913666
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
ORGANISM    Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 921)
AUTHORS     Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE       Generation of ESTs from potato callus tissue
JOURNAL     Unpublished (2003)
COMMENT     Other_ESTs: EST740324
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES            Location/Qualifiers
     source          1..921
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POCCV33"
                     /tissue_type="callus"
                     /lab_host="DH10B-TonA"
                     /clone_lib="potato callus cDNA library, normalized and
                     full-length"
                     /note="Vector: pCMVsport6.1; Site 1: EcoRI; Site_2: NotI;
                     supplier: RNA was isolated from Solanum tuberosum var.
                     Kennebec callus tissue grown on solid media."

ORIGIN
Alignment Scores:
Pred. No.:           1.47e-51        Length:         921
Score:               839.00          Matches:        185
Percent Similarity:  72.63%          Conservative:    22
Best Local Similarity: 64.91%        Mismatches:      48
Query Match:         57.82%          Indels:          30
DB:                                  Gaps:             7
```

```
Qy   194  ProAspAspSerLysSerTyr-----------------------------------ArgSerArgArgSer  206
          |||::::||||||||                                              ||:||||||||
Db   622  CGCAGTCGGAGCCCGTTCTTACTCCAAAGGAAAGAGTGTTAGTCGAAGCCGTAGCCGTAGC  681

Qy   207  ArgGlyProSerCysSerTyrSerSerIleValSerProAlaArgSerIle  226
          |||||||||   ||:||||||||||||||||||
Db   682  CGTAGCCGAAGCAAAAGT---CGTAGCAAGAGCAAAAGCAAAATCCTAAAGTTAAGTCT  738

Qy   227  SerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSer  246
          |||||||   |||||||||||::  |||||||||||||:::
Db   739  TCAAAGCGTCCAAGA-----TCTCGTTCAAGGTCTGTG--TCTTCTCGGTCTCGTTCT  789

Qy   247  GlySerLeuLeuArgAlaGlyAlaAspTrpIleSerGlnSerArgSerLysArgSerArg  266
          |||::::
Db   790  GGG-----------------------------TCAAAAGGAGCGTCCGTGTCTCAAGATCTCCA  822

Qy   267  SerArgSerAsnSerPro  272
          |||||||||||||||||
Db   823  TCAAGGTCCAGATCGCCA  840
```

RESULT 12

CG848162

LOCUS   CG848162   874 bp   DNA   linear   GSS 18-NOV-2003
DEFINITION   ZMMBb0319K18.r ZMMBb Zea mays subsp. mays genomic clone
             ZMMBb0319K18 3', genomic survey sequence.
ACCESSION   CG848162
VERSION   CG848162.1 GI:38375023
KEYWORDS   GSS.
SOURCE   Zea mays subsp. mays (maize)
ORGANISM   Zea mays subsp. mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 874)
  AUTHORS   Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
             and Wing,R.
  TITLE   Sequencing of the maize genome
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Rod Wing
             Arizona Genomics Institute
             University of Arizona
             Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
             85721-0088, USA
             Tel: 520 626 3967
             Fax: 520 621 9288
             Email: http://genome.arizona.edu
             PCR Primers
             FORWARD: T7
             BACKWARD: M13r
             Plate: 0319 row: K column: 18
             Seq primer: M13r
             Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..874
                     /organism="Zea mays subsp. mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /sub_species="mays"
                     /db_xref="taxon:4578"
                     /clone="ZMMBb0319K18"
                     /lab_host="DH10B"
                     /clone_lib="ZMMBb"
                     /note=vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Zea mays L. ssp. mays"
ORIGIN

Alignment Scores:
Pred. No.:   874
Score: 829.50
Percent Similarity: 75.09%   Matches: 180
Best Local Similarity: 65.93%   Conservative: 25
Query Match: 57.17%   Mismatches: 43
DB: Indels: 25
                                    Gaps: 8

Pred. No.:   6.79e-51   Length: 874
```

grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN

```
Alignment Scores:
Pred. No.:              3.02e-51    Length:         896
Score:                  834.50      Matches:        184
Percent Similarity:     72.73%      Conservative:    24
Best Local Similarity:  64.34%      Mismatches:      49
Query Match:            57.51%      Indels:          29
DB:                                 Gaps:             7
```

US-10-014-927-19 (1-279) x CK275124 (1-896)

```
Qy     3  SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu  22
          ||||||   |||||||||:::||||||       ||||||||||||||||||||||||
Db    22  AGTCGGTCGGAGTAGGACACTTAATGTTGGCAATCTTCCTGGTGATGTCCGTGAGCGTGAA  81

Qy    23  ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspIleLeuProPro  42
          |||||||||||:::|||||||   |||||||||||||||
Db    82  GTGGAAGATCTGTTCACAAGATGGCCCGATAGCTCATATTGAGCTGAAATTCCACCA  141

Qy    43  ArgProProGlyValAlaPheValGluPheGluAspProArgAspAlaAspAlaIle  62
          ||||||||||   ||||||||||:::
Db   142  AGACCCCCTGGTTATGTCTTTTGTTGAGTTGAAGAGGCTCGTGATGGTGAAGATGCTATT  201

Qy    63  TyrGlyArgAspGlyTyrAspPheAspGlyCysArgGlyLeuArgValGluIleAlaHisGly  82
          ||||||||||||||   |||||||||||||||
Db   202  CGTGGTCGTGATGGCTATGATTTTGATGGGACCGGTTGCGGGTTGAGCTTGCACATGGC  261

Qy    83  GlyArgArgPheSerProSerValAlaSerArgTyrSerSerSerTyrSerAla---------  99
          ||||||||||||
Db   262  GGGGGCCGTGGTAACTCATCATCAAATGATCGTTATGGTGGCGGCGGCGTGGTGGCGGC  321

Qy   100  ---------SerArgAlaProSerArgSerAspTyrArgSerValLeuValThr  114
          ||||||||||||||||||:::
Db   322  GGCCGTGGTCAGCGTGGTGGAGAGTGTCCAGACGTTCGATTTTCGAGTGTTAGTTACA  381

Qy   115  GlyLeuProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAsp  134
          ||||||||||||:::
Db   382  GGATTACCCCATTCAGCAATCGTGGCAGGATCTCAAGGATCACATGCGTCGTGCTGGGGAT  441

Qy   135  ValCysPheSerGluValPheProArgLysGlyMetSerGlyValValAspTyrSer  154
          ||||||||   ::::||||||||||:::
Db   442  GTTTGTTTCTCCCAAGTTTTCCGTGATGGGAGGTGCACCACTGGTTATAATTGATTATACA  501

Qy   155  AsnTyrAspAspMetLysTyrAlaIleArgIleLeuAspAlaThrGluPheArgAsnAla  174
          |||||||||   ||||||:::
Db   502  AACTATGATGATGAACATGCTCATCAAAAAGCTTGATGACTTCGAGTTTCGGAATGCC  561

Qy   175  PheSerSerAlaTyrIleArgValArgGluTyrGlu---SerArgSerValSerArgSer  193
          ||||||||||||||||||||:::
Db   562  TTTTCTCGTGCCACAATTCGACAATTCGAGTTAAGGAATATGACCCGAGTCGAAGCGCCAGCCGCAGC  621
```

REFERENCE   1 (bases 1 to 903)
AUTHORS   Buell,C.R., Hart,A., Ziemann,V., Karamycheva,S.A., Day,B.,
             Staskawicz,B., Jin,H. and Baker,B.
TITLE   Generation of EST sequences from Nicotiana benthamiana
JOURNAL   Unpublished (2003)
COMMENT   Other ESTs: EST761045
             Contact: Robin Buell
             The Institute for Genomic Research
             9712 Medical Center Dr, Rockville, MD 20850, USA
             Email: potato-array@tigr.org
             Clones can be requested from TIGR via potato@tigr.org
             Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES   Location/Qualifiers
     source   1..903
                     /organism="Nicotiana benthamiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:4100"
                     /clone="NBMDF19"
                     /tissue_type="abiotic and biotic stress-treated leaves,
                     callus tissue and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="Nicotiana benthamiana mixed tissue cDNA
                     library, normalized, full-length"
                     /note="Vector: pCMVsport6.1; Site 1: EcoRI; Site_2: NotI;
                     supplier: RNA was isolated from Nicotiana benthamiana
                     tissues that include callus, roots from liquid culture
                     grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
                     cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
                     challenged leaves (Pseudomonas syringae pv tomato 12 hr;
                     Xanthomonas campestris pv campestris 12 hr, 18hr;
                     Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
                     campestris pv vesicatoria 18hr). RNA was isolated from
                     these tissues and pooled in approximately equal molar
                     amounts."

ORIGIN

Alignment Scores:
Pred. No.:   7.1e-51   Length: 903
Score: 829.50   Matches: 185
Percent Similarity: 74.48%   Conservative: 28
Best Local Similarity: 64.69%   Mismatches: 44
Query Match: 57.17%   Indels: 29
DB: 14   Gaps: 9

US-10-014-927-19 (1-279) x CK298330 (1-903)

Qy   3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
          :::|||||||||||::||:||||:|||||||||||:::||
Db   79 AGTCGATCCAGTAGGACACTTTATGTTGGCAATCTTCCGGGCCGGATATTCGTGAGCGTGAA 138

Qy   23 ValGluAspLeuPheTyrLysTyrGlyTyrGlyProIleValAspIleAspLeuLysIleProPro 42
          ||||||||:||:|||||||||||||||||||||::||||:|||
Db 139 GTGGAAGATCTGTTTTACAGTATGGCCCAATAGCGCATATTGAGCTGAAAATTCCACCA 198

Qy   43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62
          |||:||||||||:::||||:::||||||||||||||||
Db 199 AGGCCCCCGGTTATGCTTTTGTGAGTTTGAAGAGTTCGTGATGCTGAAGATGCTATT 258

Qy   63 TyrGlyArgAspGlyTyrAspPheArgSerGlyCysArgLeuArgValGluIleAlaHisGly 82
          ||::||||::|||||||||||||:::|||||||||
Db 259 CGTGGTCGTGATGGCTATGATTTTGATGGGAGGTTGAGCTTGCACCATGGT 318

RESULT 13
CK298330
LOCUS   CK298330   903 bp   mRNA   linear   EST 15-DEC-2003
DEFINITION EST761044 Nicotiana benthamiana mixed tissue cDNA library,
            normalized, full-length Nicotiana benthamiana cDNA clone NBMDF19 5'
            end, mRNA sequence.
ACCESSION   CK298330
VERSION   CK298330.1 GI:39885597
KEYWORDS   EST.
SOURCE   Nicotiana benthamiana
ORGANISM   Nicotiana benthamiana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.

US-10-014-927-19 (1-279) x CG848162 (1-874)

Qy   1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
          ||||| :::|||||||||||::||:||||||||
Db 115 ATGAGCTAGCCGTTCCAGCAGGACTCTTTACGTTGGAAATCTTCCGGGTGATGTTCGGCGAG 174

Qy   21 CysGluValGluAspLeuPheTyrLysTyrGlyTyrGlyProIleValAspIleAspLeuLysIle 40
          ||||||||||||||||||||:|||||||||||||:::|||
Db 175 AGGGAAGTGGAAGATTTGTTTTATAGTATGGTCCTATAGCTCAAATTGACTTGAAGAT 234

Qy   41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
          |||||||||||||::||||:::||||||||:||
Db 235 CCACCAAGGCCTCCAGGTTATGCATTTGTTGAGTTTGAAGAAGCTCGAAGATGCTGAAGAT 294

Qy   61 AlaIleTyrGlyArgAspGlyTyrAspPheArgSerGlyCysArgLeuArgValGluIleAla 80
          |||:||||||||||||::::||||::::|||||||||
Db 295 GCCATTCGTGGTCGTGATGGATATGACTTTCGAGGGCATCGTTTACGGGGTTGAACTTGCT 354

Qy   81 HisGlyIleArgArgPheSerProSerValAspArgTyrSerSer---SerTyrSerAla 99
          ||::||||||||:::||||||||::||||::||
Db 355 CATGTGGTCGTGGGCGTCA---TCCATAGATCGTCACAGCAGTTATATGTAGTGGGGCGT 411

Qy 100 SerArgAlaProSerArgArgAspSerArgAspTyrArgValLeuValThrGlyLeuProProSer 119
          |||:::||||::||||||||:::||||||||||||||||
Db 412 GGACGTGGACCATCCAGGCGTTCGAATATCGCGTGCTAGTTACTGACCATCTTCT 471

Qy 120 AlaSerTrpGlnAspLeuLysAspHisMetSerGlyMetSerGlyValValValAspTyrSerAsnTyrArgAspAspMet 139
          ||||||||||||||||:::|||||||||::::|||||||||||||
Db 472 GCTTCATGGCAGGACCTCAGGATCACATCGTCGAGCAGGAGATGTTTGTTTTTCCCAA 531

Qy 140 ValPheProAlaArgGlyMetSerGlyValValValAspTyrSerAsnTyrArgAspAspMet 159
          ||||| :::||||||||||::::||||||||||
Db 532 GTTTTCCGTGATGGTAGTAGTGGACTACACGGAGGATTGTGGACTACACCAACTATGATGATATG 591

Qy 160 LysTyrArgAlaIleArgLysLeuAspAlaThrGlyLeuPheArgAsnAlaPheSerSerAlaTyr 179
          ||||||| :::||||||||||||::||||||||||
Db 592 AAGTATGCTATTAAGAAACTCGATGACTCTGAGTTTCGTAATGCATTTTCGGGCATAT 651

Qy 180 IleArgValArgGluTyrGluSer---ArgSerValSerArgSerProAspAspSerLys 198
          |||:::||||||||||:||:::|||||||||||||:::||
Db 652 GTTCGGGTTAAGGAATATGATTTCAGGCGGGATTCCTCTCGAAGCCCCAGTCGTTGGCCGA 711

Qy 199 SerTyr---ArgSerArgArgSerArgGlyTyrProSerCysSerTyrSerLysSer 217
          |||::: ||:::|||||||
Db 712 TCTCTCTCCAAGAAGCCAGAAGCCGAGCCGA-------------AGTAGAAGC 750

Qy 218 ArgSerValSerProAlaArgSerIleSerSerProArgSerArgProLeuSerArgSerArg 237
          ||||| |||:::||||||||||::::||
Db 751 CGAAGCAGAAGCCGTGGTCGGAGCTAT-----------AGCCGAAGCAGAGCCGCAGCAAG 801

Qy 238 SerLeuTyrSerSerValSerArgSerGlySerLeuLeuArgAlaGlyAlaGlyAspTrpProIleSer 257
          ||||| |||::||||||||||::||
Db 802 TCTCCAAAGGCAAAGCCTTCCGCGAAG---------------------------TCA 831

Qy 258 GlnSerSerLysSerArgSer---ArgSerArgSer 269
          :::||| |||||||::::|||||
Db 832 CCTGAAAAGTCTANGATCAGGTCTCTCCTCGTTCCCGCTCT 870

grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Alignment Scores:
Pred. No.:   7.4e-51   Length: 930
Score: 829.50   Matches: 185
Percent Similarity: 74.48%   Conservative: 28
Best Local Similarity: 64.69%   Mismatches: 44
Query Match: 57.17%   Indels: 29
DB: Gaps: 9

US-10-014-927-19 (1-279) x CK285308 (1-930)

```
Qy    3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
          |||     |||||  ||  | :::  ||| |||| :::  |||  |||||||
Db   49 AGTCGATCGAGTAGGACACTTTATCGTGCAATCTTCCGGGCGATATTCGTCGAGGCTGAA 108

Qy   23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuAspLysIleProPro 42
          ||  |||   |||  ||    |||||  ||::: |||  ||| |||  ||| |||||
Db  109 GTGGAAGATCTGTTTTACAAGTATGGCCCAATAGCGCATATTGAGCTGAAAATTCCACCA 168

Qy   43 ArgProProGlyTyrAlaPheValGluAspProArgArgAlaAspAspAlaIle 62
          |||  |||  ||     |||||||  |||  |||  |||  ||| |||
Db  169 AGGCCCCCTGGTTATGCTTTTGTTGAGGAGTTTGAAGGAGTTCGTGATGCTGAAGATGCTATT 228

Qy   63 TyrGlyArgAspAspProTyrAspPheAspGlyCysArgLeuArgValGluIleAlaAlaHisGly 82
          |||   |||   ||| | |:::   ||| |||||  |||  |||  ||||  || :::
Db  229 CGTGGTCGTGATGGCTATGATTTTGATGGGCACCGTCTGAGGGTTGAGCTTGCACATGGT 288

Qy   83 GlyArgArgPheSerProSerValAspArgTyrSerSerSerAlaSerArg--- 101
          ||:::  ||  |||  | ||  || :::  |||  |||  || :::
Db  289 GGGGCGTGGTCAGTCA---TCAAACGATCGTTATGGTGGTGGTGGAGGTCGTGGTCAACGT 345

Qy  102 ---AlaProSerArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
             :::  |||  |||  ||| :::  |||  |||  |||  ||| ||| :::
Db  346 GGTGGAGTGTCCAGACGTTCGATTATGCGTGAAAATTACAGGATTACCCCATTCAGCG 405

Qy  121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
          ||| ||| |||  |||  |||   |||  ||  ||| :::  |||  |||  |||
Db  406 TCATGGCAGGATCTCAAGGATCCAAGTCCGTCGAGCTGGGAGTGTCTGTTTCTCCAAGTT 465

Qy  141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrArgAspMetLys 160
          ||  |||  |||  ||| :::  ||| |||  |||  |||  |||  |||  |||
Db  466 TTCCGTGAACGAGGGGGACTACTGGTATTGTGGATTATACAAACTACGATGACATGAAA 525

Qy  161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
          |||  ||  |||  |||  |||     |||  |||  ::: |||  |||
Db  526 TATGCTATCAAAAAGCTTGATGCTACTGAGTTTCGGAATGCTTTTCGTTCAACAATA 585

Qy  181 ArgValArgGluTyrGlu---SerArgSerValSerArgSerProAspSerLysSer 199
          |||  |||  |||:::     |||  |||  ||::: |||  |||
Db  586 CGGGTTAGGGAATATGACCACAGCCCGCAGCTGCAGCCGCAGC--- 627

Qy  200 TyrArgSerArgSerArgArgGlyProSerGlyProSerCysSerTyrSer---SerLysSerArg 218
          |||  |||  |||  |||  |||  |||  |||  |:::   :::  |||  |||
Db  628 ---CGCAGCCGTTCTTACTCCAGAGAAAAGACGGGCAGTAAAAGCCGTAGCCGAAGTAGA 684

Qy  219 SerValSerProAlaArgSer--------IleSerProArgSerArgPro 232
          |||  |||  ||| |||  ::: 
Db  685 AGTGTCAGTCGTAGCAGGAGCAAAAGCAAATCTCCAAAGTTAAATCTTCAAGGGCGTCA 744

Qy  233 LeuSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuLeuArgAla 252
          ||| :::  |||  |||  |||     |||  |||
Db  745 AGATCTCGTTCAAGGTCGTC---TCTTCTCCGGTCCCGTTCTGGA------- 786

Qy  253 GlyAspTrpIleSerArgGlnSerArgSerLysArgSerArgSerArgSerAsnSerPro 272
```

Qy 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrArgAspMetLys 160
Db 496 TTCCGTGAACGAGGGGGACTACTGGTATTGTGGATTATACAAACTACGATGACATGAAA 555

Qy 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 556 TATGCTATCAAAAAGCTTGATGACTCTGAGTTTCGGAATGCCTTTTCGTTCAACAATA 615

Qy 181 ArgValArgGluTyrGlu---SerArgSerValSerArgSerProAspSerLysSer 199
Db 616 CGGGTTAGGGAATATGACCACAGCCCGCAGCTGCAGCCGCAGC------ 657

Qy 200 TyrArgSerArgSerArgGlySerArgGlyProSerTyrSer---SerLysSerArg 218
Db 658 --CGCAGCCGTTCTTACTCCAGAGAGAGGAAAGACCCTAGCCGAAGTAGA 714

Qy 219 SerValSerProAlaArgSer--------IleSerProArgSerArgPro 232
Db 715 AGTGTCAGTCGTAGCAGGAGCAAAAGCAAATCCCAAAAGTAAATCTTCAAGGGCGTCA 774

Qy 233 LeuSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuLeuArgAla 252
Db 775 AGATCTCGTTCAAGGTCCGTC---TCTTCTCCGGTCCCGTTCTGGA------ 816

Qy 253 GlyAspTrpIleSerArgGlnSerArgSerLysSerArgSerArgSerArgSerAsnSerPro 272
Db 817 ------------TCAAAGGGGCGTTCTATCTCCAGATCTCCATCAAGGTCCAGATCGCCA 864

Qy 273 ValSerProValIleSer 278
Db 865 --GCACCATCGTAAGT 879

RESULT 14
CK285308
LOCUS   CK285308   930 bp   mRNA   linear   EST 15-DEC-2003
DEFINITION EST748030 Nicotiana benthamiana mixed tissue cDNA library,
            normalized, full-length Nicotiana benthamiana cDNA clone NBMAU11 5'
            end, mRNA sequence.
ACCESSION   CK285308
VERSION   CK285308.1   GI:39859739
KEYWORDS   EST.
SOURCE   Nicotiana benthamiana
  ORGANISM Nicotiana benthamiana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1 (bases 1 to 930)
  AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
            Staskawicz,B., Jin,H. and Baker,B.
  TITLE   Generation of EST sequences from Nicotiana benthamiana
  JOURNAL   Unpublished (2003)
COMMENT   Other ESTs: EST748031
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES   Location/Qualifiers
     source   1..930
                     /organism="Nicotiana benthamiana"
                     /mol_type="mRNA"
                     /db_xref="taxon:4100"
                     /clone="NBMAU11"
                     /tissue_type="abiotic and biotic stress-treated leaves,
                     callus tissue and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="Nicotiana benthamiana mixed tissue cDNA
                     library, normalized, full-length"
                     /note="vector: pCMVsport6.1; Site 1: EcoRI; Site 2: NotI;
                     supplier: RNA was isolated from Nicotiana benthamiana
                     tissues that include callus, roots from liquid culture

```
Db    787 ----------TCAAAAGGGCGGTTCTATCTCAAGATCTCCATCAAGGTTCCAGATCGCCA  834
                     ||||::     ||||||  ||||||  |||||||||    |||||||
Qy    273 ValSerProValIleSer                                          278
                 :::|||       ::|||
Db    835 ---GCACCATCTGTAAGT                                          849
```

```
RESULT 15
CK265450
LOCUS       CK265450            884 bp    mRNA         linear   EST 12-DEC-2003
DEFINITION  EST711528 potato abiotic stress cDNA library Solanum tuberosum cDNA
            clone POABY79 5' end, mRNA sequence.
ACCESSION   CK265450
VERSION     CK265450.1  GI:3982428
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 884)
  AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
  TITLE     Generation of ESTs from abiotic stressed potato tissue
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: EST711529
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from TIGR via potato@tigr.org
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1..884
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POABY79"
                     /tissue_type="abiotic stress treated leaf and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="potato abiotic stress cDNA library"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: Solanum tuberosum var. Kennebec plants were
                     grown from cuttings on a 16hr light/8 hr dark cycle at 25
                     C for 3-4 weeks. Abiotic stress conditions were applied to
                     four separate sets of plants. Set 1 involved saturation of
                     the soil with 150 mM NaCl and tissues were harvested at
                     following application of the salt stress (leaves: 2hr,
                     6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
                     Set 2 were grown under the standard conditions and then
                     were water stressed by withdrawal of further watering
                     applications. Drought stressed plants were harvested after
                     cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
                     and 5d). Set 3 were grown under the standard conditions
                     and then were cold stressed by placement at 4 C. Cold
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d. Set 4 were grown under the standard conditions and
                     then were heat stressed by placement at 35 C. Heat
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d and 4d and heat-stressed roots were harvested at 6 hr,
                     12 hr, 1 d, and 4d. RNA was isolated from all tissues and
                     equal RNA from each tissue and stress was pooled to
                     construct the cDNA library. RNA sample."
ORIGIN
```

Alignment Scores:
```
Pred. No.:          7.5e-51        Length:          884
Score:              829.00         Matches:         182
Percent Similarity: 71.78%         Conservative:     24
Best Local Similarity: 63.41%      Mismatches:       47
Query Match:        57.13%         Indels:           34
DB:                 14             Gaps:              7
```

US-10-014-927-19 (1-279) x CK265450 (1-884)

```
Qy     3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu    22
           |||||:::       ::::|||||||||||||   :::|||||||
Db    23 AGTCGGTGGAGTAGGACACTTATGTTGGCAATCTTCCTGGTGATGTTCGTGAGCGTGAA    82

Qy    23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProPro    42
           |||||||||||||||||::|||    :::|||||||||||||
Db    83 GTGGAAGATCTGTTCACACAGTATGGCCCGATAGTCATATTGAACTGAAAATTCCACCA   142

Qy    43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAspAlaIle    62
           ||||||||||||::|||   :::|||   ::::|||||||||||
Db   143 AGACCCCCTGGTTATGCTTTGTTGAGTTTGAGGATCCGCGATGCTGAAGATGCTATT     202

Qy    63 TyrGlyValArgAspGlyTyrAspPheAspGlyCysSerArgLeuArgValGlyIleAlaHisGly  82
           ::::||||||||||:::|||    :::|||||||||||
Db   203 CGTTGGTCGTGATGGCTATGATTTTGATGGGCACCACCGTCTGCGGGTTGAGCTTGCACATGGC   262

Qy    83 GlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAla-----------  99
           |||||||||| :::|||  |||||||
Db   263 GGGGGTGGTGTAACTCATCATCAAATGATCGTTATGGTGGCGGCGGTGGTGGTGGTGGTGGC   322

Qy   100 ------------SerArgAlaProSerArgSerAspTyrArgValLeuVal   113
                        :::||||||||
Db   323 GGCGGCCGTGGTCAGCGTGGTGGAGGAGTGTCCAGACCGTTCTGATTTTCGAGTGTTAGTT   382

Qy   114 ThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGly   133
           |||||||||||  :::||||||||||
Db   383 ACAGGATTACCCCATTCAGCATCGTGGCAGGATCTCAAGGATCACATGCTGTGCTGGG     442

Qy   134 AspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyValValAspTyr   153
           |||||||||::|||   :::|||||||||
Db   443 GATGTTTGTTTCTCCCAAGTTTCCGTGATGGGAGTGGCACCACTGGTATAATTGATTAT   502

Qy   154 SerAsnTyrAspAspMetLysTyrAlaIleArgLeuAspAlaThrGluPheArgAsn     173
           ::|||   |||||||||||||::::|||||||||
Db   503 ACAAACTATGATGACAATGAAATATGCTATCAAAAAGCTTGATGACTCGAGTTTCGGAAT   562

Qy   174 AlaPheSerSerAlaTyrIleArgValArgGluTyrGlu---SerArgSerValSerArg   192
           ||||||::::|||   :::|||||||||
Db   563 GCCTTTTCTCGTGCGACAATTCGAGTTAAGAAATATGACCCGCAGTCGAAGCCGCAGCCGC   622

Qy   193 SerProAspAspSerLysSerTyr---------------------ArgSerArgSerArg   205
           |||::::|||
Db   623 AGCCGCAGTCGGAGCCGGTTCTTATTCCAAAGGAAAGAGTTGTTAGTCGAAGCCGTAGCCGA   682

Qy   206 SerArgGlyProSerCysSerValSerSerSerValSerProAlaArgSer   225
           ||||||||||||::::|||
Db   683 AGTAGAAGT----------CGTAGCAAGAGCAAAAGCAAATCCTAAAGTTAAG   727

Qy   226 IleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArg   245
           ::|||||||||||||
Db   728 TCTTCGTTCAAGGCTCAAGA------TCTTCGTTCAAGGTCTGTG---TCTTCGGTCTCGT   778

Qy   246 SerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgLysSerArgSerArgSer   265
           |:::||
Db   779 TCTGGG-------------------------------TCAAAAAGGACGTCCTGTCTCCAGATCT   811

Qy   266 ArgSerArgSerAsnSerPro   272
           ||||||
Db   812 CCATCAAGGTCCAGATCGCCA   832
```

Search completed: March 26, 2004, 20:24:40
Job time : 2837 secs

(1026)

```
RESULT 5
ATH131214
LOCUS           ATH131214               5164 bp    DNA     linear   PLN 07-MAY-1999
DEFINITION      Arabidopsis thaliana srp30 gene, exons 1-12.
ACCESSION       AJ131214
VERSION         AJ131214.1  GI:4775269
KEYWORDS        SF2/ASF-like splicing modulator; srp30 gene.
SOURCE          Arabidopsis thaliana (thale cress)
ORGANISM        Arabidopsis thaliana
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE       1
AUTHORS         Lopato,S., Kalyna,M., Dorner,S., Kobayashi,R., Krainer,A.R. and
                Barta,A.
TITLE           atSRp30, one of two SF2/ASF-like proteins from Arabidopsis
                thaliana, regulates splicing of specific plant genes
JOURNAL         Genes Dev. 13 (8), 987-1001 (1999)
MEDLINE         99234087
PUBMED          10215626
REFERENCE       2  (bases 1 to 5164)
AUTHORS         Barta,A.
TITLE           Direct Submission
JOURNAL         Submitted (02-DEC-1998) Barta A., Institute for Biochemistry,
                University of Vienna, Dr. Bohrgasse 9/3,, A-1030 Vienna, AUSTRIA
FEATURES             Location/Qualifiers
     source          1..5164
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
     mRNA            join(1806..2014,2138..2203,2290..2364,2444..2546,
                     2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
                     3501..3578,4515..4572,4844..5125)
     exon            1806..2014
                     /number=1
     gene            1928..4858
                     /gene="srp30"
```

```
CDS       join(1928..2014,2138..2203,2290..2364,2444..2546,
          2632..2678,2786..2843,2968..3017,3161,3291..3403,
          3501..3578,4515..4572,4844..4858)
          /gene="srp30"
          /codon_start=1
          /product="SF2/ASF-like splicing modulator Srp30"
          /protein_id="CAB42557.1"
          /db_xref="GI:4775270"
          /db_xref="GOA:Q9XFR5"
          /db_xref="SPTREMBL:Q9XFR5"
          /translation="MSSRWNRTIYVGNLPGDIRKCEVEDLFYKYGPIVDLKIPPRP
          PGYAFVEFEDPRDADDAIYGRDGYDFDGCRLRVEIAHGGRRFSPSVDRYSSSYSASRA
          PSRRSDYRVLVTGLPPSASWQDLKDHMRKAGDVCFSEVFPDRKGMSGVVDYSNYDDMK
          YAIRKLDATEPRNAFSSAYIRVREYESRSVSRSPDDSKSYRSRSRSRGPSCSYSSKSR
          SVSPARSISPRSRPLSRSRLYSSVSRSQSRSKSRSRSRSNSPVSPVISG"

CDS       join(1928..2014,2138..2203,2290..2364,2444..2546,
          2632..2678,2786..2843,2968..3017,3161,3291..3403,
          3501..3578,4178..4214)
          /gene="srp30"
          /note="alternative exon 10a"
          /codon_start=1
          /product="SF2/ASF-like splicing modulator Srp30, variant
          1"
          /protein_id="CAB42558.1"
          /db_xref="GI:4775271"
          /db_xref="GOA:Q9XFR6"
          /db_xref="SPTREMBL:Q9XFR6"
          /translation="MSSRWNRTIYVGNLPGDIRKCEVEDLFYKYGPIVDLKIPPRP
          PGYAFVEFEDPRDADDAIYGRDGYDFDGCRLRVEIAHGGRRFSPSVDRYSSSYSASRA
          PSRRSDYRVLVTGLPPSASWQDLKDHMRKAGDVCFSEVFPDRKGMSGVVDYSNYDDMK
          YAIRKLDATEPRNAFSSAYIRVREYESRSVSRSPDDSKSYRSRSRSRGPSCSYSSKSR
          SVSPARSISPRSRPLSRSRLYSSVSRSGSLLRAGDWI"

intron    2015..2137
          /gene="srp30"
          /number=1

exon      2138..2203
          /gene="srp30"
          /number=2

intron    2204..2289
          /gene="srp30"
          /number=2

exon      2290..2364
          /gene="srp30"
          /number=3

intron    2365..2443
          /gene="srp30"
          /number=3

exon      2444..2546
          /gene="srp30"
          /number=4

intron    2547..2631
          /gene="srp30"
          /number=4

exon      2632..2678
          /gene="srp30"
          /number=5

intron    2679..2785
          /gene="srp30"
          /number=5

exon      2786..2843
          /gene="srp30"
          /number=6

intron    2844..2967
          /gene="srp30"
          /number=6

exon      2968..3017
          /gene="srp30"
          /number=7

intron    3018..3160
          /gene="srp30"
          /number=7

exon      3161..3217
          /gene="srp30"
```

```
intron    3218..3290
          /gene="srp30"
          /number=8

exon      3291..3403
          /gene="srp30"
          /number=9

intron    3404..3500
          /gene="srp30"
          /number=9

exon      3501..3578
          /gene="srp30"
          /number=10

intron    3579..4514
          /gene="srp30"
          /number=10

exon      4178..4572
          /gene="srp30"
          /note="alternative exon"
          /number=11

exon      4178..4340
          /gene="srp30"
          /note="alternative exon 10a"

exon      4515..4572
          /gene="srp30"
          /number=11

intron    4573..4843
          /gene="srp30"
          /number=11

exon      4844..5125
          /gene="srp30"
          /number=12

ORIGIN

Alignment Scores:
Pred. No.:                          Length:        5164
Score:             5.92e-39         Matches:       242
Percent Similarity: 891.00          Conservative:  0
Best Local Similarity: 43.60%       Mismatches:    3
Query Match:       43.60%           Indels:        313
DB:                61.41%           Gaps:          9
                   8

US-10-014-927-19 (1-279) x ATH131214 (1-5164)

Qy    1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys   20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1928 ATGAGTAGCCGATGGAATCGTACGATCTACGTTGGGAATTTGCCTGGAGATATTCGCAAG 1987

Qy   21 CysGluValGluAspLeuPheTyrLys----------------------------------   29
        |||||||||||||||||||||||||
Db 1988 TGTGAGGTTGAAGATCCTTCTACAAGGTTTGAAAATTCCTCTTTTCTCTGATAAAAAT 2047

Qy   29 ------------------------------------------------------------   29
Db 2048 TGAATTCATTATGACTAGTTGGGTTCATAAATTTGCAATTCGTCTTGCTGAGACAATT 2107

Qy   30 --------------------------------------TyrGlyProIleValAspIleAspLeuLys   39
                                              |||||||||||||||||||
Db 2108 TAAATCGACTCTTATGTATATTGTTCAGTATGGACCAATTGTGGACATTGATTTGAAG 2167

Qy   40 IleProProArgProProGlyTyrAlaPheValGlu--------------   51
        ||||||||||||||||||||||||||
Db 2168 ATTCCACCGAGACCTCCTGGTTATGCCTTTGTCGA-GGTATTGATCAAGTACAAATTT 2226

Qy   51 --------------------------------------------------   51
Db 2227 GTTTTTTCTTCCTTCTTGTATAGTATAGGCTAAGATGACTAAGAGTTGTTATTCGTGG 2286

Qy   52 ---PheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPhe   70
           |||||||||||||||||||||||||||||
Db 2287 CAGTTTGAAGATCCTCGTGATGCAGACGATCAATTTATGACGTGATGGTTATGATTTT 2346

Qy   71 AspGlyCysArgLeuArg------------------------   76
        ||||||||||||||||||
```

```
Qy  219 ------------SerValSerProAlaArgSerIleSerProArgSerArgProLeuS  234
Db 3485 TCTAAATGCAATGCAGTGTGTCACCTGCTAGATCCATTTCCCGCGTTCACGGCGCCCTTA  3544

Qy  234 erArgSerArgSerLeuTyrSerSerValSerArg  245
Db 3545 GTCGGTTCTCGCTTGCTATACAGCTCGTGTCTCAAGG  3579
```

```
Db 2347 GATGGGTGTCGACTTCGGGTTAGTAAACGCATGATGAAAGCTAGCTTAATTTCTGTAAT  2406
Qy   77 -----------------------------------ValGluIleAlaHisGlyGlyAr   84
Db 2407 TTCTTGTAAAGGTGTTATCTTTGTGTGATGTTTTAGGTTGAGATTGCACATGGTGGTCG  2466
Qy   84 gArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAlaProSe  104
Db 2467 TAGATTTTCACCATCAGTTGATAGTAGTACAGCAGCAGCCGTGCGAGCCGGTGCACCTTC  2526
Qy  104 rArgArgSerAspTyrArg-------  110
Db 2527 AAGAGCTCTGACTACCGCGGTTGTGAGTCTTCTCGATTGTGTTATTGGTGTGTGT  2586
Qy  111 --------------------ValLeuValThrGlyL  116
Db 2587 AAAATTTTATATTTGAAAACTCATTTTACTACCTAAACATGTAGTGCTTGTGACCGGAT  2646
Qy  116 euProProSerAlaSerTrpGlnAspLeuLys---------  126
Db 2647 TACCGCCTTCTGCTTCGTGGCAGGACCCTAA-GGTAAGGGACACTATATAGTCTTTTCT  2705
Qy  126 ------------------  126
Db 2706 CTGAATGTTGGTTCTCTATATCATGTTTGATTTATCTCTTTTCTGAAATGATGTTATT  2765
Qy  127 -------AspHisMetArgLysAlaGlyAspValCysPheSerGluV  140
Db 2766 TGCTATTTACGGGTGATTAGGATCACATGCGCAAAGCTGGAGATGTCTGCTTCTCTGAAG  2825
Qy  140 alPheProAspArg-Lys-------------  145
Db 2826 TTTTCCCTGACCGTAAAGGTGAGTTGACATTCGATAGTTTGGATAAGCTTTTTGATTGAT  2885
Qy  145 ------------------  145
Db 2886 GTGTTAGTAAATTAGTCTTTGTGAAGAGAATAGGTGTTAAGCATCTGAACTGCTAAACT  2945
Qy  146 -----GlyMetSerGlyValValAspTyrSerAsnTyrAspAsp  158
Db 2946 CACATTCAGTATTTCTTTGTAGGCATGTCTGGGGTTGTGGATTATAGCAACTATGATGAT  3005
Qy  159 MetLysTyrAla-------  162
Db 3006 ATGAAGTACGC-AGTAAGTTTTATATCTTTGCAACGCAAATGTTCCTGGACTTATGCCTT  3064
Qy  162 ------------------  162
Db 3065 AGACTGCTTTTGTTTGTTTCATGATAGTATACCGGAGCTGAATTTATCTTCCTGGAGGCCAGTGTTGG  3124
Qy  163 -------------IleArgLysLeuAspAlaThrGlu  170
Db 3125 ATCTTTGATGTTCCCTTAAAATTTTGATGTGACAGATAAGGAAACTTAAAAGCTATGAAGCAGGAGTTCGGAG  3184
Qy  171 PheArgAsnAlaPheSerSerAlaTyrIleArg-------  181
Db 3185 TTTTCGAAATGCTTTTCTCTCAGTGCTTATATACCGGTGTATGTTGATTGTTTCTTTGTTCTTTGATTTT  3244
Qy  182 --------------ValArgGluTyrGl  186
Db 3245 GTTAAGCATAAGTGGTGATATGGATGAGTCATCTCGAATTCACTCTTGTTCAGGTGAGGGGGAATATGA  3304
Qy  186 uSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSe  206
Db 3305 GTCGGAGGAGTGTGAGCTGCAAGCCCAGATGATTCTAAAAGCTATAGAAGCAGGAGAGGAGTCGGAG  3364
Qy  206 rArgGlyProSerCysSerTyrSerSerLysSerArg-------  218
Db 3365 CCGTGCGTCCCAAGCTGCTAGCTATAGTAGCAAGAGCAGAGGAGTCTTTTTTTTTTTTTT  3424
Qy  218 ------------------  218
Db 3425 TCATAAACCTAAGACATATAAGGGATTTTATTGTAACTTATTATTGAAATTAACTAACT  3484
```